WARWICK
THE UNIVERSITY OF WARWICK

# Pushing the limits of two-dimensional mass spectrometry

**Dissertation submitted in partial fulfilment for the Degree of Doctor of Philosophy**

Federico Floris, MSc AMRSC

November 2017

Department of Chemistry

# Contents

**ABSTRACT:** Two-dimensional Fourier transform ion cyclotron resonance mass spectrometry (2D FT-ICR MS) is a data independent tandem mass spectrometry technique that allows direct correlation between precursor and fragment ions without the necessity of any sort of isolation prior to fragmentation. Two-dimensional mass spectrometry (2D MS) experiments were conceived in the 1980s from inspiration of two-dimensional nuclear magnetic resonance (2D-NMR), but the development of the technique was stopped because of the insufficient computational capabilities of that time. In 2010, with the progress in computer technology and the improvements in FT-ICR MS instrumentation and methods, the interest in developing 2D MS was renewed. Since then, 2D FT-ICR MS showed to be a valuable platform for the analysis of small molecules, macromolecules such as peptides and glycopeptides, and complex mixtures deriving from tryptic digestion of proteins of increasing sizes. All the proteomics studies performed with 2D MS used the bottom-up (BUP) approach.

The main goal of this Ph.D. research work was to develop two-dimensional mass spectrometry for the challenges of contemporary mass spectrometry, therefore pushing the limits of the technique towards the most interesting, cutting-edge research grounds, by optimising the technique, developing 2D MS analysis methods, and finally constructing improved instrumentation act to better perform 2D MS experiments.

On this purpose, the rising importance of top-down proteomics (TDP) led to the development of 2D FT-ICR MS for the analysis of intact proteins. To realise this first project, 2D MS was optimised for TDP analysis using calmodulin (CaM) as a model, and infrared multiphotondissociation (IRMPD) as fragmentation technique. The study compared the two-dimensional BUP and TDP analysis of CaM with standard one-dimensional tandem mass spectrometry, and showed that the use of 2D MS allowed to obtain comparable cleavage coverages, with a consistent saving in sample and time. The study affirmed the suitability of 2D MS for top-down proteomics, still leaving space for more developments allowing its optimisation. The total cleavage coverage of CaM was later implemented with the use of BUP 2D electron-capture dissociation (ECD) MS.

In order to improve the cleavage coverages obtained in top-down two-dimensional mass spectrometry, another technique was developed, called MS/2D MS. The technique makes use of an extra fragmentation step before two-dimensional mass spectrometry, achieved through external isolation of a charge state of interest, and collisional fragmentation in a hexapole-based collision cell before the ions are sent to the ICR-cell. The output of MS/2D MS is a single 2D mass spectrum with information equivalent to $MS^3$ about the entire mixture. The technique has been successfully developed using ubiquitin (Ubi) and calmodulin as models, and showed new uses of

the features exclusive of 2D MS, such as the use of precursor ion scans to assign the protein terminus of many precursors ions at once, and the analysis of internal fragments formed by the first dissociation. Furthermore, as long as extra fragmentation steps can be performed, a deeper investigation of the analytes can be achieved, in an experiment collectively called $MS^n$/2D MS.

After the development of 2D MS for TDP, another method was developed, implementing the use of two-dimensional mass spectrometry for another type of complex mixtures: polymers. An extensive 2D MS study was performed on a simple mixture of homopolymers such as PMMA, and on two complex mixtures commonly used as drug excipients, polysorbate 80 and TPGS. The study demonstrated again that single 2D MS experiments give results equivalent to multiple MS/MS experiments targeted at the different precursors. Furthermore, it allowed the high resolution tandem mass spectrometry of polysorbate 80, representing the first kind of study of that type because of the challenges of isolating single analytes from the polymer mixture without the use of a previous separation, such as chromatography. The use of neutral-loss lines was particularly useful in this study, allowing to discriminate between oleated, linoleated, and non esterified species. Finally, 2D mass spectra of polymers showed characteristic diagonal lines, specific for each polymeric mixture, representing some sort of fingerprint of the mixture.

The last part of the thesis is dedicated to the optimisation of 2D FT-ICR MS for ECD MS, the application of the technique to another class of compounds, glycans and glycan-derivatives, and finally to the development of a more capacious ICR-cell.

The research performed in this thesis showed that 2D FT-ICR MS is a viable data-independent mass spectrometry technique for top-down proteomics and polymer analysis; that the technique is fast, and it possesses unique features that improve ion assignments, such as precursor ion scans and neutral-loss lines. Finally, it opens for two-dimensional mass spectrometry the path to glycomics and to an improved 2D MS analysis of even more complex mixtures, thanks to the development of a novel ICR-cell.

# List of Figures

# List of Tables

# Nomenclature

$\omega$ frequency

$\omega_c$ cyclotron frequency

$\omega_m$ magnetron frequency

$\omega_z$ z-oscillation frequency

$m/z$ mass-to-charge ratio

$M_n$ number average molecular weight

$M_w$ weight average molecular weight

ABC ammonium bicarbonate

ACN acetonitrile

AI-ECD activated-ion electron-capture dissociation

AP atmospheric pressure

APCI atmospheric pressure chemical ionisation

APPI atmospheric pressure photoionisation

BUP bottom-up proteomics

CAD collisionally-activated dissociation

CaM calmodulin

CHEF correlated harmonic excitation field

CI chemical ionisation

CID collision-induced dissociation, aka CAD

CRS charge-reduced species

CSV comma-separated values

CytC cytochrome c

ECD electron-capture dissociation

EDD electron detachment dissociation

EI electron ionisation

ESI electrospray ionisation

ETD electron transfer dissociation

ETnoD electron transfer no-dissociation

ExD electron-based dissociation

FFT fast Fourier transform

FID free-induction decay

FT Fourier transform

FT-ICR Fourier transform ion cyclotron resonance

FWHM full width at half maximum

GC gas chromatography

GC-MS gas chromatography mass spectrometry

hECD hot electon-capture dissociation

HPLC high pressure liquid chromatography

ICI in-cell isolation

ICR ion cyclotron resonance

IR-ECD infrared-activated electron-capture dissociation

ISD in-source dissociation

LC liquid chromatography

LC-MS liquid chromatography mass spectrometry

MAI matrix-assisted ionisation

MALDI matrix-assisted laser desorption/ionisation

MeOH methanol

MMA mass measurement accuracy

MS mass spectrometry

MS/MS tandem mass spectrometry

MS/MS/MS two-dissociations tandem mass spectrometry, aka MS3

MSn tandem mass spectrometry with n stages

nESI nano-electrospray ionisation

nETD negative electron transfer dissociation

niECD negative-ion electon-capture dissociation

NMR nuclear magnetic resonance

OVAT one variable at a time

PEG poly(ethylene glycol)

PMMA poly(methyl) methacrylate

ppm part per million

PTM post-translational modification

RF radio frequency

RP resolving power

SORI sustained off-resonance irradiation

SORI-CAD sustained off-resonance irradiation CAD, also SORI-CID

SubP substance P

SVD single value decomposition

TDP top-down proteomics

TPGS D-α-Tocopherol polyethylene glycol 1000 succinate

Ubi ubiquitin

UHV ultra-high vacuum

UV/vis ultraviolet/visible spectrophotometry

UVPD ultraviolet photodissociation

# Copyright

# Preface

Before you lies the thesis *Pushing the limits of two-dimensional mass spectrometry*, fruit of my efforts to develop a rediscovered technique to face the challenges of contemporary mass spectrometry. It has been written to fulfill the requirements to obtain the crown of Doctor of Philosophy in Chemistry, and it represents the work of three years of research spent behind chemicals, computers, and - running or dismantled - instrumentation.

The journey began with my acceptance of a Ph.D. offer that made me leave my job for a multinational company, move from the lovely city of Liverpool to - ehm - Coventry (look up "send to Coventry"), and say goodbye to the small party of friends I built in nine months of permanence in the U.K. only to find out that the project I had applied for was not available anymore, but that some budget was available to develop an old, niche, mass spectrometry technique. With the positivity that had always characterised me before the start of this Ph.D., I accepted the new challenge, and started learning on my own the principles of electronics, programming, Fourier transform ion cyclotron resonance mass spectrometry, and finally its underdeveloped two-dimensional variant. About what happened afterwards, I could write for ages about the sense of awe and agony that underpinned my doctoral life span, but a line is already enough. This is my biggest academic achievement so far, and I prefer highlighting much more: the passion that drove me from failed experiments to less failed ones; the curiosity that guided me from a discover to the other; the pride in my research, which pushed me to spread the word, and find many different audiences for my little contributions to Science, in the U.K, in Europe, and in the U.S. I prefer seeing the last three years as an adventure, an epic quest to revive the sacred spirit of 2D MS and be recognised as a King - ehm - Doctor. That said, every quest has its difficulties, traps, mid-bosses, and final boss, and after three publications as first author, one as a co-author, and the submission of a fourth manuscript and a (spell) book chapter, I am writing this scroll to earn my victory.

I would like to thank the companions that assisted me in my doctoral adventure, starting from my supervisor and the people from his research group that actively contributed to the happy ending of this story. I would also like to thank a long list of external collaborators, all present in my publications, and the partners, friends and acquaintances that were part of the journey.

Finally, a Story is nothing without a reader. Therefore, a special "thank you" goes

to whoever goes through these pages, with my wish for you to enjoy the reading and that it will ignite in you the spirit of adventure.

Federico Floris

Sìnnai, Sardinia, Italy, August 2017

# Declaration and inclusion of material

This thesis is submitted to the University of Warwick in support of my application for the degree of Doctor of Philosophy. It has been composed by myself and has not been submitted in any previous application for any degree.

The work presented (including data generated and data analysis) was carried out by the author except in the cases outlined below:

- Section 6.2.1 - The samples used for the 2D MS analysis of glycans and glycopeptides where prepared in collaboration with Anne Sophie Gleinich, Division of Biomedical Sciences, Warwick Medical School, University of Warwick. *A.S.Gleinich@warwick.ac.uk
- Section 6.3 - The open cylindrical cell initial design was a project by Prof Peter B. O'Connor and Dr Rebecca H. Wills (Department of Chemistry, University of Warwick), who also commisioned the production of the cell plates. *p.oconnor@warwick.ac.uk
- Section 6.3.1 - The cell plates where produced by the Department of Engineering at the University of Cambridge. The magnetic susceptibility measurements were performed by Dr Martin Lees, Department of Physics, University of Warwick. *M.R.Lees@warwick.ac.uk
- Section 6.3.3 - The printed circuit board was designed and produced in collaboration with Dr Alex W. Colburn, Department of Chemistry, University of Warwick. *A.W.Colburn@warwick.ac.uk

# 1. Introduction



## 1.1. Mass spectrometry

Mass spectrometry is a qualitative and quantitative analytical technique for the identification of molecules in mixtures of different complexities. The technique is based on the detection of the mass-to-charge ratio ($m/z$) of the analytes in the gas phase, used then to assign their molecular structures. Mass spectrometry (MS) was born more than 100 years ago, and since its conception it has been greatly developed, to become finally an important investigation tool in every laboratory where analytical chemistry is practiced. Particularly in the last decades, development of new mass analysers and detectors, and hyphenation of different mass analysers with the result of hybrid instruments brought MS to be an essential platform in the fields of proteomics, metabolomics, glycomics, and other "omics".

### 1.1.1. Principles of mass spectrometry

Each mass spectrometric instrument can be described according to the scheme in Figure 1.1. In order to be manipulated, analysed, and detected, the analytes need to be ionised. Charging of the particles occurs in the ionisation source (Fig. 1.1.a), with consequent production of ions based on the technique and on the analytes' ionisation efficiencies. Different ionisation sources produce, in fact, different kinds of ions, *e.g.* by addiction of a proton or a metal ion, interaction with electrons, detachment of protons, and so on. Consequently, ions produced can be positive or negative. Ionisation techniques are divided in two main groups, based on the ability to generate a molecular ion from an analyte and keep it intact, or to fragment it during ionisation, thus generating fragment ions. Ionisation techniques that generate ion fragments in the ionisation source are defined as "hard", opposite to the ones that generate only or mainly the molecular ion, called "soft" ionisation techniques. Both the hard and soft ionisation techniques find their applications in routine analysis and research, with preference for the latter, and some examples are elaborated below.

Once the ions are formed, they can be manipulated through electromagnetic fields, in order to be selected and guided towards the mass analyser (Fig. 1.1.b), which discriminates them based on their mass-to-charge ratio. The ions are subsequently detected by the mass detector (Fig. 1.1.c), which records a signal (generally an electric signal) proportional to their abundance. The signal is finally acquired and processed in a data system (Fig. 1.1.d), and a mass spectrum is produced. A mass spectrum is a plot of the abundances of the detected ions against their $m/z$.

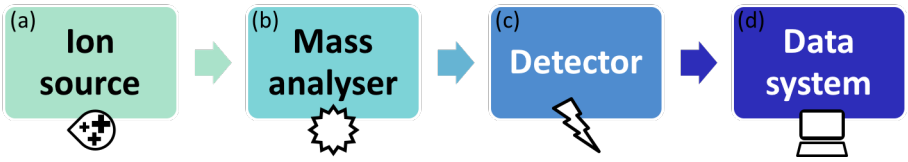



Figure 1.1.: Scheme of an MS instrument. The sample is introduced in the ionisation source, where the analytes are ionised. The resulting ions are then transferred to a mass analyser, which discriminates between them based on their $m/z$, and subsequently to a mass detector, the part of the instrument that allows to observe the signal. The signal is then acquired and processed in a data system, and finally a mass spectrum is produced.

A wide range of mass spectrometers exists on the market, based on different mass analysers and detectors and/or their hybridisation; on their portability, or specific uses. Existent mass analysers include but are not limited to quadrupoles, magnetic sectors, ion traps, time-of-flight (ToF), and Fourier-transform based analysers such as FT-ToF, FT-ICR MS, and Orbitrap. All of the mass spectrometry work for this thesis has been performed with Fourier transform ion cyclotron resonance (FT-ICR) MS technology, with other methods referred to as needed.

## 1.2. Fourier transform ion cyclotron resonance mass spectrometry

Fourier transform ion cyclotron resonance mass spectrometry (FT-ICR MS) is the mass spectrometric technique with the highest performances in terms of resolving power and mass accuracy. It is based on the motion of ions in a high magnetic field, and the detection of their resulting ion cyclotron frequency, inversely proportional to their mass-to-charge ratio. For its high resolution and mass accuracy, FT-ICR MS has become an important platform in the major fields interfacing physical chemistry and biology, such as metabolomics, glycomics, and proteomics, but also retains the leading position in the investigation of complex mixtures, such as those deriving from crude oils and other petroleum-based products, in a relatively new category field called “petroleomics” [2]. The technique was invented in 1974 by Comisarow

and Marshall [3, 4]by implementing FT to existing ICR technology, and today it is practiced worldwide.

### 1.2.1. Principles of FT-ICR MS

FT-ICR MS follows the same principles explained in Figure 1.1 for every mass spectrometry technique. The analytes are introduced in the ionisation source as solids, liquids or gases, and then eventually moved into the gas phase during ionisation. The gas-phase ions are subsequently transferred, separated according to their $m/z$, and finally detected. In FT-ICR MS, the mass detector and mass analyser coincide, embodying the core of the instrument, the ICR-cell. The cell is a vessel constituted by different electrodes and enclosed in a superconducting magnet, used to trap the ions in motion.

Charged particles moving in a high magnetic field with a component of their velocity perpendicular to the magnetic field axis experience a force known as Lorentz force. For ions that move only perpendicularly to the magnetic field axis, assuming that no electric field is present, the Lorentz force acting on the charged particles is given by

$$F_L = qvB \tag{1.1}$$

Where $q$ is the charge on the particle (in Coulombs, C), $v$ is its velocity (m/s) and $B$ is the magnetic field strength (Tesla, T). The total charge possessed by an ion can be described as the number of electronic charges in the ion, $z$, times the electric charge per electron $e$ (Coulombs, C): $q = ze$.

Ions moving in different directions will follow a precession along the magnetic field axis, with a frequency inversely proportional to their mass-to-charge ratio. This frequency is known as cyclotron frequency, and it is described by the equations

$$f_c = \frac{qB}{2\pi m} \tag{1.2}$$

$$\omega_c = \frac{qB}{m} \tag{1.3}$$

Where $m$ is the mass of the ion (kg), and $f_c$ and $\omega_c$ are respectively the cyclotron frequencies in Hertz, Hz, and radiants per second, rad/s . As it can be seen from equations 1.2 and 1.3, the cyclotron frequency is inversely proportional to the $m/z$ of the ions, allowing their discrimination and subsequent identification/assignment.

In order to be detected, ions moving inside a high magnetic field need to be trapped. The ICR-cell traps the ions, and allows for their mass analysis and detection. Detection of the ions inside the cell can be described through a brief explanation of the

ICR-cell structure. A schematic representation of a cubical FT-ICR analyser cell is shown in Figure 1.2, bottom. The cell is aligned with the bore of a superconducting magnet in a way that the magnetic field resultant is coaxial with the centre of the cube ($z$-axis). The electrodes that constitute the cell are divided in excitation and detection plates, opposite to each other. At the end of the cell, two trapping plates are positioned perpendicularly in respect to the other electrodes. Inside the ICR-cell, the ions move around the magnetic field axis in packets according to the experienced Lorentz Force: detection of the ion packets occurs as they move across the detection plates (Fig. 1.2, top right). Their passage causes the charges on the plates to move within the detection circuit to counteract the proximity of the ions (an electron transit known as image current). This change in electrical potential can be measured as a function of time, giving the raw data in the time domain, called transient signal. Detection is a non-destructive process, as the ions are repeatedly detected as they pass close to the detection plates. The magnitude of the signal is given by the proximity to the ions to the detection plates, independently of the magnetic field. Ions inside the ICR-cell at thermal energy, however, do not have a sufficiently high orbital radius to be detected. In order to be detectable, the ions must be excited to a higher orbital radius through the use of a radio-frequency potential (RF) applied at the excitation plates (Figure 1.2, top left). This RF potential has to be in resonance with the ions cyclotron frequencies so that the ions are excited coherently. The RF potential can be swept with a linear increase through the interested frequencies range (RF "chirp") [4] or applied through a technique known as stored-waveform inverse Fourier transform (SWIFT) [5], in which a rectangular excitation frequency spectrum of the wanted frequencies is transformed through inverse Fourier Transform to a waveform that is applied to the excitation plates.

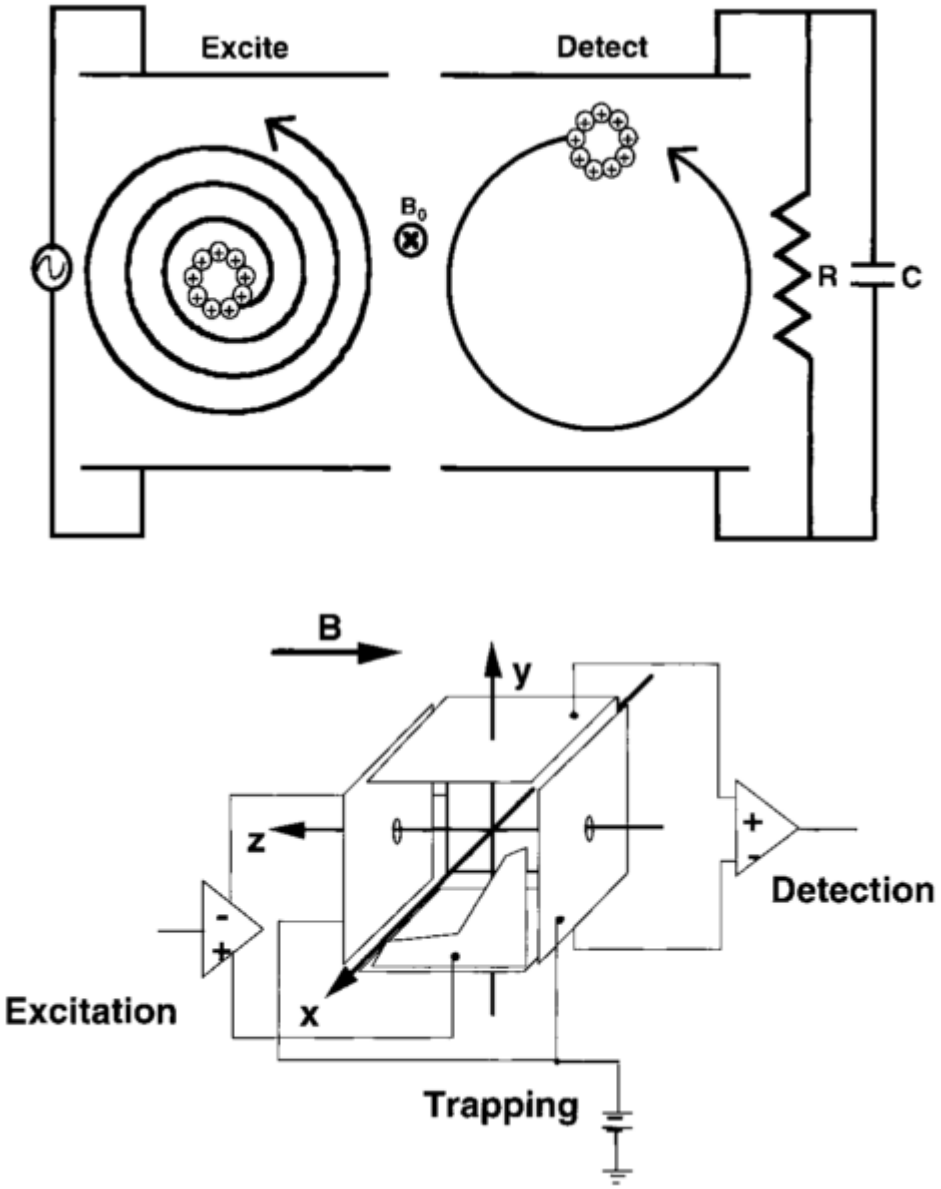



Figure 1.2.: Representation of excitation and detection (top) inside an ICR-cell (bottom). Top: behaviour of the ions inside an ICR-cell when a d.c. potential is applied to the excitation electrodes (top left) and when the ions are close to the detection electrodes (top right). Bottom: representation of a cubical ICR-cell, and the circuitry used to apply the excitation pulses, and subsequently detect the ion charges. Ions enter the ICR-cell along the $z$-axis, and are trapped by the voltage applied to the trapping electrodes, perpendicular to such axis. Reproduced from Marshall *et al.* [6].

The trapping plates at the end of the cell are used to restrain the ions inside the vessel. A low potential is applied to these plates, usually of the order of 1 V. Although low, the application of this potential ensures that the electrostatic field inside the cell is not zero, and that the cyclotron motion does not occur along the exact centre

of the cell. This configuration gives rise to the magnetron motion, bringing the ions to move according to three components in total (Figure 1.3): cyclotron motion, magnetron motion, and oscillation between the trapping plates (parallel to the $z$-axis). The magnetron motion causes the ions to precess along higher orbits around the magnetic field axis, basically acting as the cyclotron motion, but at a smaller frequencies. The magnetron frequency depends on geometry of the cell (dimensions, design), the trapping potential, and the magnetic field strength. It can be expressed by the equations:

$$f_m = \frac{V_{trap}\alpha}{2\pi B a^2} \tag{1.4}$$

$$\omega_m = \frac{V_{trap}\alpha}{B a^2} \tag{1.5}$$

Where $f_m$ and $\omega_m$ represent the magnetron frequency respectively in Hz and rad/s, $V_{trap}$ is the trapping potential (in Volts, V), $a$ is the distance between the trapping plates, and $\alpha$ is a factor dependent on the geometry of the cell. The equations show the independence of the magnetron motion from the $m/z$ of the charged particles inside the ICR-cell, determining the impossibility of any analytical use for such phenomenon. Ions otherwise stable subjected to the magnetron motion continue to expand their precession along the magnetic field axis, and they are eventually lost. Currently the only effective way of returning ions to the centre of an ICR cell despite magnetron motion is the use of quadrupolar axialisation. Such technique allows to convert magnetron motion to unperturbed cyclotron motion (Equation 1.3), allowing re-measurement of stored ion packets. Theoretically, the possibility to measure stored ion packets an infinite number of times enables an "infinite sensitivity" measurement, where the same ions are repeatedly detected and returned to the ICR cell, possibly down to single-ion detection. Furthermore, this feature allows to accumulate many scans of the same species without introduction of new ion packets and subsequent usage of sample.

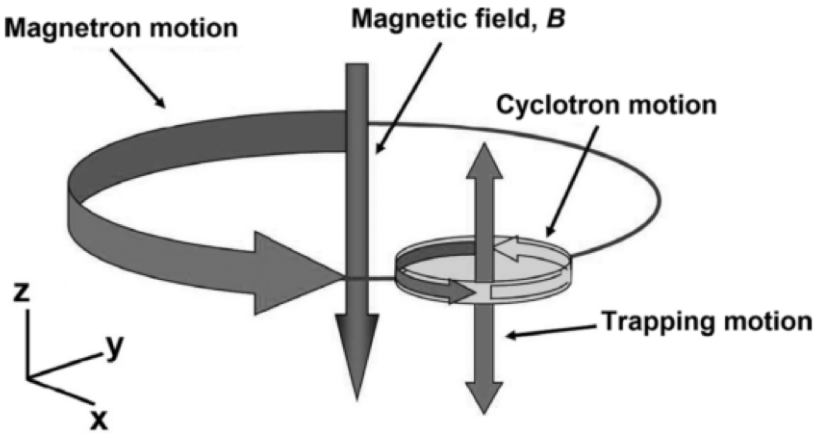



Figure 1.3.: Ion motions inside an ICR-cell. The magnetic field is along the $z$-axis. Courtesy of Bruker Daltonics, Bremen, Germany.

Due to the presence of different motions, the measured cyclotron frequency is reduced according to

$$\omega_c^- = \frac{\omega_c}{2} - \sqrt{\left(\frac{\omega_c}{2}\right)^2 - \frac{\omega_z^2}{2}} \tag{1.6}$$

where $\omega_z$ is the "trapping" oscillation frequency, given by the equation $\omega_z = \sqrt{\frac{2qV_{trap}\alpha}{ma^2}}$.

The cyclotron frequency of the ions within the magnetic field is independent of their kinetic energy, difficult to control and cause of major variance in mass analysers that use it to discriminate different mass-to-charge ratios, such as ToF-MS analysers [6]. However, the motion of ions discussed above assumes an ideal vacuum, which is unfortunately unobtainable. Collisions between ions and background gas molecules can perturb ion motion causing the cyclotron frequency to decrease to an undetectable level. Excessive collisions with background gas can quickly lead to loss of ions from the ICR cell, preventing measurement. As shown in Figure 1.4, ions in motion inside a magnetic field that collide with gas molecules inside the ICR-cell see a rapid decrease of their cyclotron motion in favour of an increase of their magnetron motion, which slowly pushes them towards the electrodes. It is for this reason that FT-ICR MS is performed in conditions of ultrahigh vacuum (UHV), despite the drawbacks of its setup.

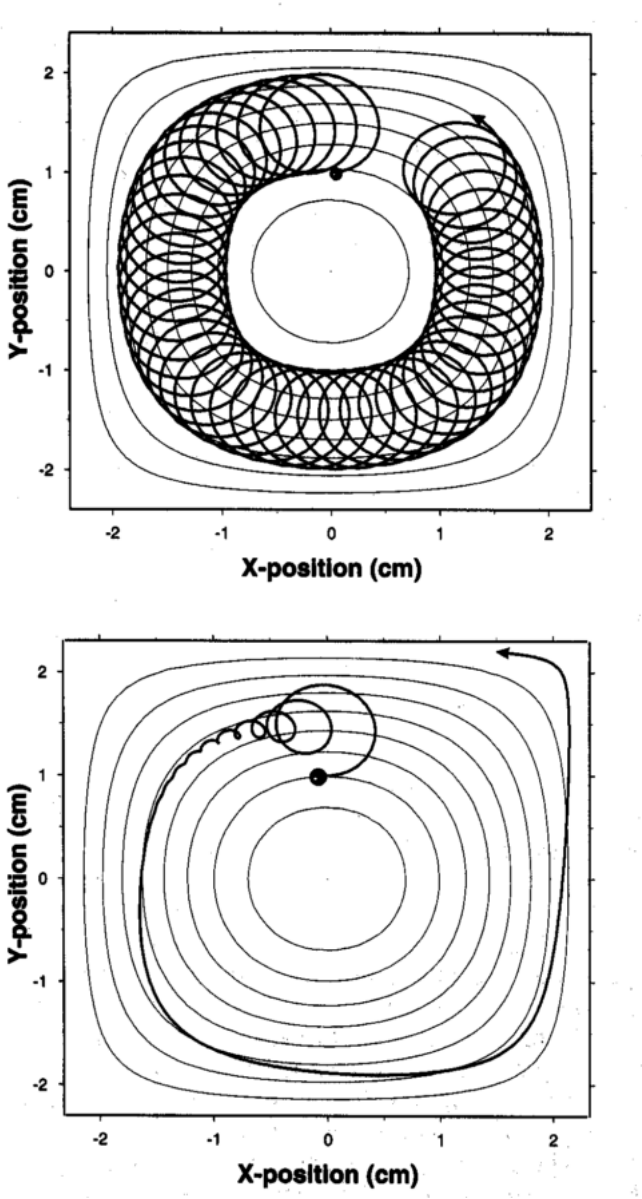



Figure 1.4.: Ion cyclotron motion and magnetron motion inside a cubical cell. The magnetic field is perpendicular to the page, and the oscillation along the $z$-axis is not shown. Top: The magnetron motion, at slower frequencies, shows a higher radius than the cyclotron frequency. Bottom: Ions moving inside the ICR-cell at their cyclotron and magnetron frequencies change behaviour in the presence of gas molecules. The cyclotron motion decreases quickly, while the magnetron motion increases, bringing the ions closer to the cell electrodes, until they are neutralised. Adapted from Amster *et al.* [7]

Finally, in order to extract data from the information obtained through the transient, a mathematical process known as Fourier Transform (FT) is applied. The algorithm transforms the time-domain data as a function of the cyclotron frequencies. A spectrum of the frequencies is obtained, which is calibrated and converted into a mass

spectrum, using the relation between the cyclotron frequencies and the masses of the ions illustrated by equations 1.2 and 1.3 (Figure 1.5).

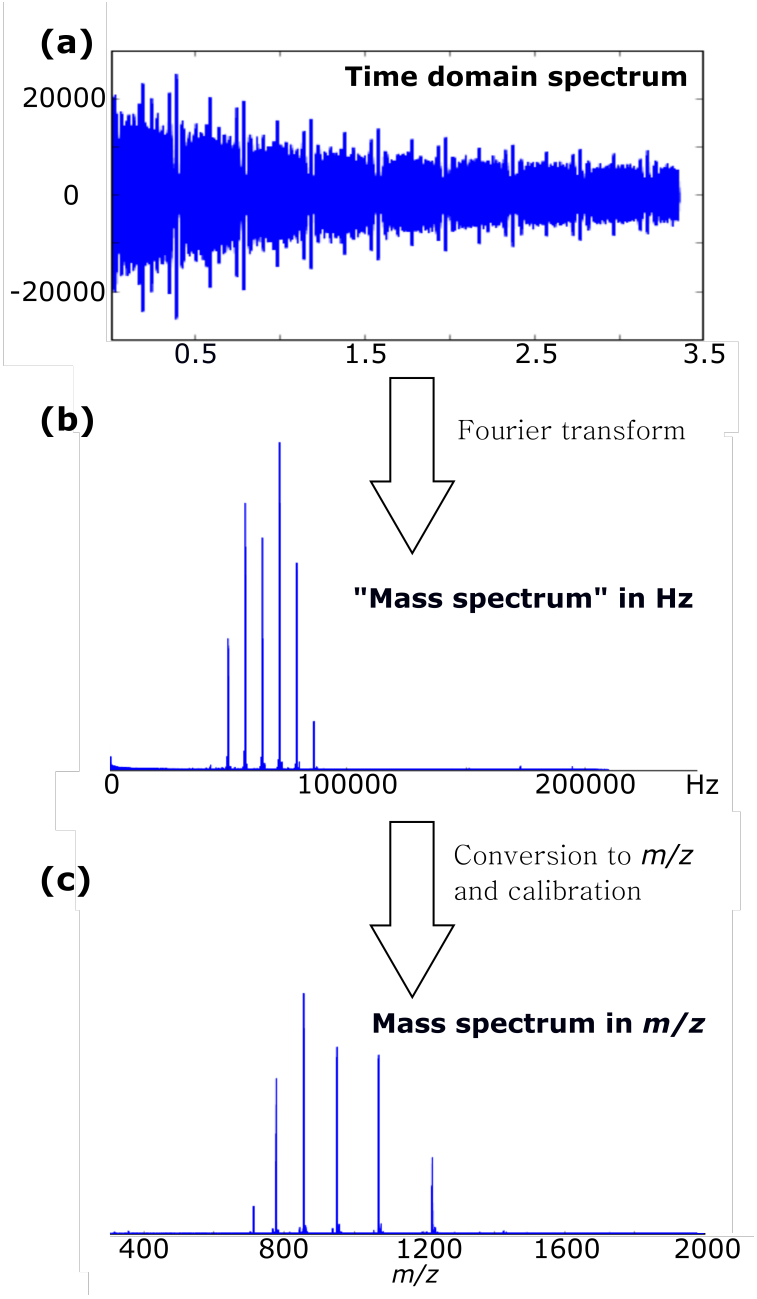



Figure 1.5.: Conversion of acquired FT-ICR MS data from the domain of time to a mass spectrum. The transient signal (a), representing the image current acquired as a function of time, is converted into a spectrum of frequencies (Hz) through a Fourier transform (b). The spectrum of frequencies is subsequently calibrated and converted into a mass spectrum (c).

### 1.2.2. Performance parameters

The parameters used universally to describe the performance of a mass spectrometry technique are mass accuracy and resolving power, and FT-ICR MS excels in both of them.

Mass measurement accuracy (MMA, abbreviated commonly in "mass accuracy") represents the efficiency of the correlation between an accurate mass, the experimentally determined mass of an ion measured to determine the elemental formula of the ion, and its exact mass, *i.e.* the calculated mass of an ion whose elemental formula, isotopic composition and charge state are known [8, 9]. The exact mass is the theoretical mass. MMA is defined by the equation $MMA = \frac{m_{obs} - m_{the}}{m_{the}}$ where $m_{obs}$ is the observed, accurate mass, and $m_{the}$ is the exact (or theoretical) mass, in Dalton. Due to the high performances in FT-ICR MS, such parameter is indicated in parts per million (ppm), corresponding to $10^{-6}$ Da. Therefore, the equation of MMA for FT-ICR MS becomes:

$$MMA = \frac{m_{obs} - m_{the}}{m_{the}} \cdot 10^6 \tag{1.7}$$

For the vast majority of mass spectrometers on the market, the absolute MMA (or AMMA) reached is in the order of tens of ppm, while for FT-ICR MS its value is indicated at around 1 ppm or lower [6].

In mass spectrometry, mass resolution represents the ability to separate two narrow peaks in the mass spectrum, and it is defined by considering two peaks of equal height and distance $\Delta m$ as

$$Resolution = \frac{m}{\Delta m} \tag{1.8}$$

where $m$ is the mass of the second peak and the distance between the two different peaks can be calculated at 0.5%, 5%, or 50% from the apex of the lower intensity peak by convention. The two peaks have to be separated by a valley not higher than 10% of their height. Resolution can be defined for single peaks as well. In that case, $\Delta m$ represents the width of the peak at a given peak height percentage, as a function of its mass. The former way to express resolution in mass spectrometry is called "10% valley definition", and the latter is called "peak width definition". The first is often associated with magnetic-sector instruments, while the second is more typical of time-of-flight or FT-MS instruments [10]. High values of resolution indicate a better performance of the instrument or the technique. In order to use resolution as defining parameter, several values of resolutions at different masses are usually indicated.

In FT-ICR MS, and generally in all types of Fourier transform spectroscopy, the term *resolution* is attributed to the width of a peak at a certain mass, usually at

50% of its height. This definition is called "full width at half maximum", FWHM. The parameter that can be used to define the mass resolution achieved by an instrument or a technique is the resolving power (RP), defined as

$$RP = \frac{m}{\Delta m} \tag{1.9}$$

Where $\Delta m$ is the resolution defined as FWHM. In FT-ICR MS, high values of resolving power determine the performance of the technique, and they are required to achieve ultra-high resolution. Resolving power is important to resolve close spaces between peaks, as the ones given by complex mixtures or multiply charged ions, and the possibility to use high-resolution instruments can make the difference in the discrimination of the species of interest. An example of this is reported by Feng *et al.* for a case where the high resolution of a 9.4 T FT-ICR mass spectrometer allowed a nitrogen-containing fullerene derivative to be identified as it was possible to resolve $^{12}C_{59}$ $^{14}N^+$ (exact mass 722.002525) and $^{12}C_{58}$ $^{13}C_2{}^+$ (exact mass 722.006161) [11]. Without the usage of such instrumentation, the presence of the nitrogen-containing species could not have been confirmed, as only one signal might have been observed, associated with the only-carbon fullerene ion.

Time-of-flight mass spectrometers usually reach resolving powers of about 20000, and magnetic-sector instruments can approach 100000. Contemporary FT-MS instruments can routinely reach RPs of hundreds of thousands, with the highest performance offered by FT-ICR MS, routinely reaching about 500000 in broadband mode (the standard experimental conditions) and a few millions in heterodyne mode [12]. Heterodyne or narrowband mode is a set of experimental conditions in which a very narrow range of $m/z$ is analysed at very high resolution (see Section 1.5.2.1).

In Fourier Transform mass spectrometry, spectra are acquired with a dataset size decided by the user in the form of data points. The number of data points determines the length of the transient, and the rate at which they are used depends on the sampling frequency. The sampling frequency is determined by the highest cyclotron frequency (*i.e.* the lowest $m/z$) set in the mass spectrum, and must be equal to at least twice that parameter, according to the Shannon-Nyqvist theorem. According to such theorem, the transient length, or acquisition time $T_{acq}$, depends on the dataset size and the sampling frequency as in the following equation

$$T_{acq} = \frac{N}{S} \tag{1.10}$$

Where $T_{acq}$ is the acquisition time in seconds, $N$ the dataset size and $S$ the sampling frequency (Hz). The resolving power is therefore dependent on the acquisition time, but it is inversely proportional to $m/z$. Increasing the low mass cut-off improves the resolution/resolving power by allowing for more data points to be used for acquisition,

provided that the ions can maintain their motion coherence inside the ICR-cell.

Finally, both mass accuracy and resolving power increase with the magnetic field strength, as shown in Figure 1.6 for different performance parameters. The advantages of having a high magnetic field push the development of bigger and better magnets, explaining why magnet design is an area of focus in FT-ICR MS research.

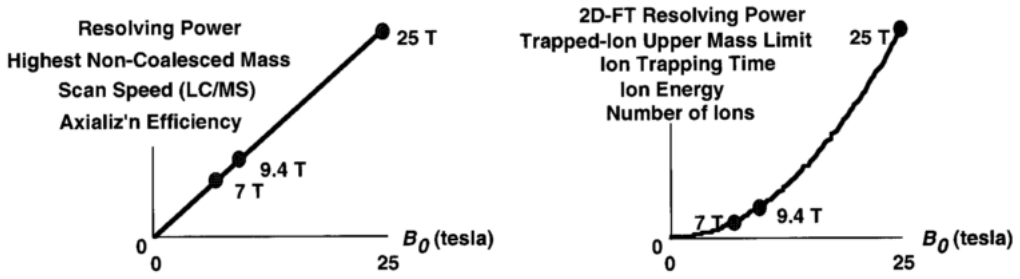



Figure 1.6.: Dependencies of different FT-ICR MS parameters on the magnetic field strength. Reproduced from Marshall *et al.* [6]

### 1.2.3. Signal acquisition

FT-ICR mass spectrometers are equipped with several independent components whose parameters can be adjusted in order to acquire the best spectra of the MS field. Users can manipulate their samples in order to improve the ionisation efficiency, the transfer of intact ions through the different parts of the instrument (skimmers, funnels, ion guides, collision or reaction cells, ion traps, flight tube, and finally the ICR-cell), their fragmentation (see "Tandem mass spectrometry" in Section 1.6), and their detection. The individual work done by the user to find the best parameters for signal acquisition, by changing the voltages in the various lenses or similar parameters and monitoring the signal strength, is called "tuning".

The main visual parameters used to tune the acquisition method are the signal strength and the peak shape of the analytes in the mixture. In contrast with the more common experimental procedure of maximising the signal strength, the high resolution of FT-ICR MS allows for the shape of the shown peaks, which should be between a Lorentzian and a Gaussian bell, to play an important role in signal acquisition. A perfect shape is an index of good motion coherence inside the ICR-cell. On the other hand, distorted peak shapes (or peak splitting) may indicate different perturbances given by the inability of the ICR-cell to contain the ions at the set parameters, ion loss, or other kinds of perturbances given by the the presence of a large number of ion packets inside the ICR-cell. Such problems are given by the interaction of the ion packet clouds, which undergo Coulombic repulsion. These interactions are of different forms, and are referred as space-charge effects [13, 14, 15, 16]. Space-charge effects reduce the achievable resolving power and mass accuracy,

and users should be aware of them in order to minimise them. When ion population inside the ICR-cell is excessive, space-charge effects can lead to a phenomenon called ion coalescence, for which two peaks of different species or isotopes of the same species converge into an observable one (Figure 1.7) [17]. By minimizing the trapping potentials, thus allowing more space for the ions to move along the $z$-axis, and carefully regulating the number of ions inside the ICR-cell, ion coalescence and other space-charge effects can be minimised. Ions inside the ICR-cell can be controlled by regulating the in-source accumulation of ions and the potential used to generate them, and by controlling the accumulation times of external ion traps (such as a hexapole) before introducing such ions in the cell.

Each ICR-cell has its own tuning parameters, and for cylindrical cells such as the Infinity cell [18] the excitation of ions at high cyclotron radii allows to increase the available space. However, for the same cell, confining the analytes to the center of the $z$-axis by increasing the trapping potential allowed to achieve ultra-high resolution in narrowband mode, previously referred to as heterodyne mode [12].

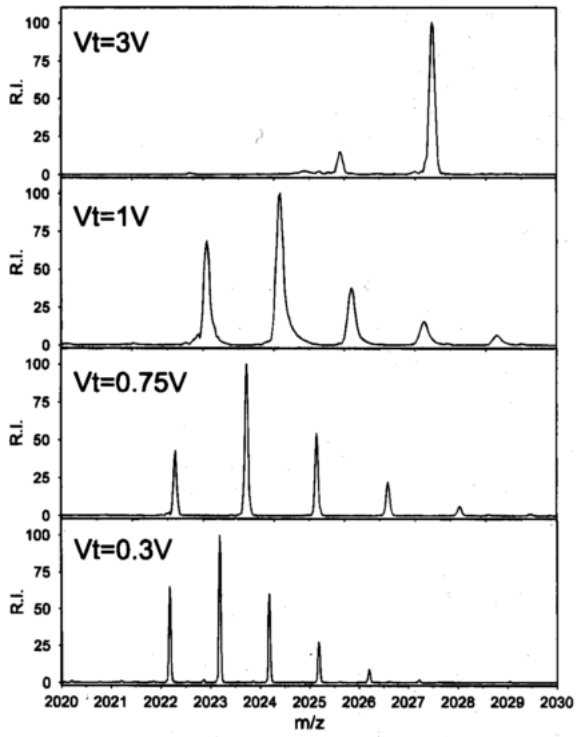



Figure 1.7.: Effect of the trapping voltages on ions of poly(ethylene glycol) (PEG) inside an ICR-cell. The application of a $V_{trap}$ of 3 V causes the immediate coalescence of the different isotopologues of PEG and the shift in frequencies, giving unexpected masses for the molecule under analysis. By diminishing the value of $V_{trap}$, gradual improvements can be observed, with the comparison of more peaks, and the increase in mass accuracy (the cyclotron frequencies shift back to acceptable mass values, situation visible by observing the space in mass between the isotopologues) and resolving power (narrower peaks). The mass spectra are reported at the relative intensities of the observed ions. Reproduced from Amster *et al.* [7]

## 1.3. Ion sources and ionisation techniques

FT-ICR MS has been coupled with almost all the existing ion sources to improve the range of analysable samples, from the simplest and most common electron ionisation (EI) and chemical ionisation (CI) [19], to niche techniques such as fast atom bombardment (FAB), atmospheric pressure photoionisation (APPI) [20, 21], atmospheric pressure chemical ionisation (APCI) [22, 23], and metastable atom bombardment [24]. Nowadays, the ionisation techniques that are mostly coupled with FT-ICR MS are matrix-assisted laser desorption/ionisation (MALDI) and electrospray ionisation (ESI). Both the techniques are labelled as "soft", as they minimise the formation of fragment ions during ionisation, with the formation of only or mainly the molecular ion. The absence of in-source fragmentation is particularly important for complex

mixtures, whose original composition, already complicated by a significant amount of often overlapping isotopic distributions, has to be analysed without the interference of fragment ions.

### 1.3.1. Matrix-assisted laser desorption/ionisation (MALDI)

Matrix-assisted laser desorption/ionisation (MALDI) is a soft technique invented in the 1990's [25], which yielded a Nobel prize to one of its developers. MALDI makes use of solid samples enclosed in a matrix, which is given the task to solvate and activate the analytes. The technique is generally performed in an under-vacuum source, even if different groups were able to achieve MALDI at pressures close to atmospheric pressure (AP) [26, 27, 28]. In MALDI, ionisation occurs by laser activation of the matrix, which absorbs the UV light undergoing photoionisation itself. The ionised matrix molecules subsequently activate the analytes by proton transfer, generating positive ions when proton adducts occur, or negative ions when a proton is detached from the sample. Other salts in the matrix can be ionised and provide adducts for the formation of ions of the molecules of interest. The process is shown in Figure 1.8. Because of the necessary laser activation, MALDI is not a continuous source, but it generates ions in pulses, dictated by the laser pulse frequency.

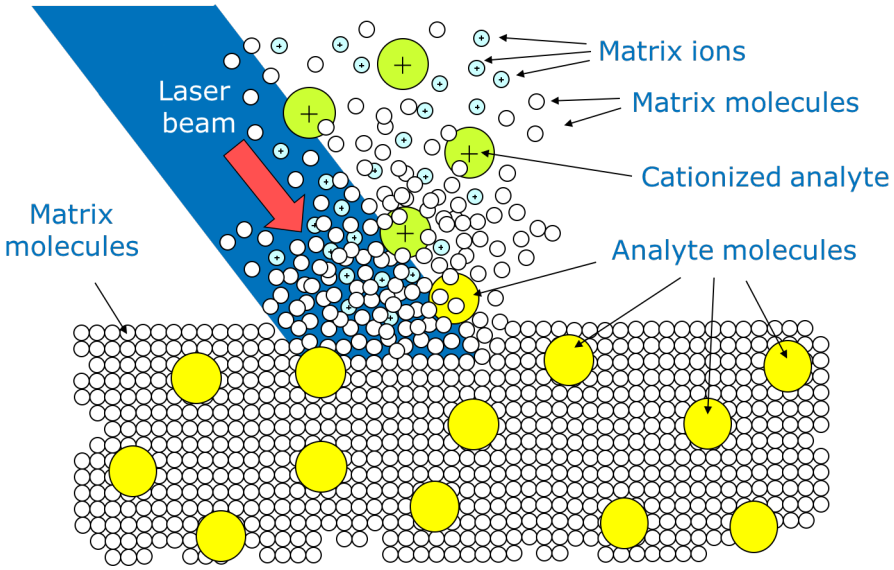



Figure 1.8.: Process of ionisation with MALDI, in this case generating positive ions. The UV-laser beam ablates the matrix surface, with consequent photoionisation of its molecules. The matrix ions transfer a proton to the analyte molecules, which become cations, and can be introduced to the core of the instrument by manipulation via electric fields. Courtesy of Bruker Daltonics, Bremen, Germany.

After the development of MALDI, intensive studies to find the best matrices for the technique were performed, with preference for molecules containing one or more

chromophores in order to absorb the UV photons necessary for their activation. Common molecules such as 2,5-Dihydroxybenzoic acid can be used as matrixes for MALDI, as long as they can absorb photons provided by the UV laser activation. Some examples are shown in Figure 1.9.

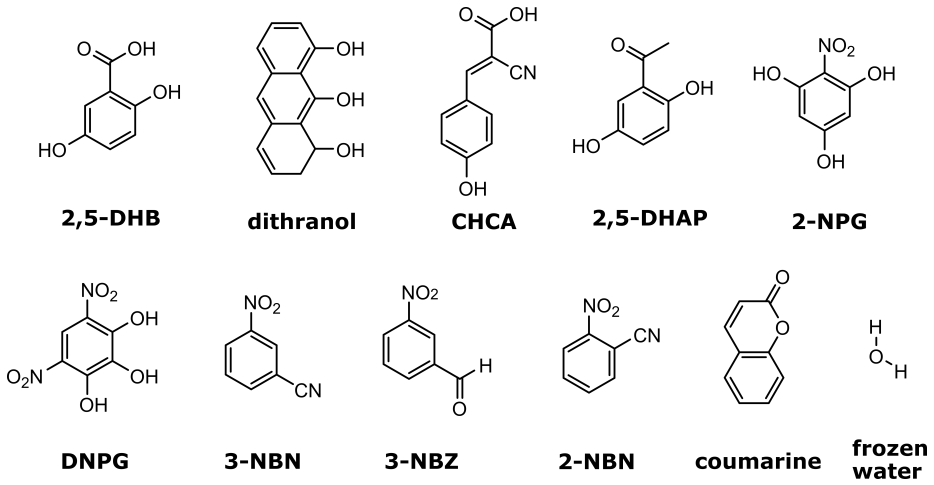



Figure 1.9.: Structures of the most common compounds used as MALDI matrices.

MALDI is very easy to use as long as a good sample preparation is performed, and it can be used for a wide range of molecules. The technique produces singly-charged ions in the vast majority of cases, a characteristic that constitutes one of its main advantages because it does not overcomplicate complex mixtures with the presence of the same species at multiple charges, but also one of its disadvantages since it limits the detection of high-mass macromolecules that can be over the upper mass limit. Furthermore, the ions produced by MALDI are generally "hotter" and more difficult to manipulate inside the mass spectrometer, meaning that a higher ion cooling by soft collisions with background gas during ion transfer and a more careful setting of the trapping plates are necessary. Finally, photoionisation of the matrix generates ions that are detected together with the molecules of interest, complicating the resulting mass spectra in the low-mass region. This situation, called matrix effect, hinders the assignment of species of interest at low $m/z$. The production of mainly singly-charged, higher-energy, ions, and the matrix effect are the reasons why ESI is often preferred to MALDI. However, MALDI is an effective ionisation technique, which demonstrated its utility in the application of MALDI-MS of small molecules, as well as biomolecules, and especially polymers, for which the generation of singly-charged polymeric distributions at high $m/z$ helps in the determination of their characteristics, such as their average molecular weights, polydispersity, and functional end groups (See Chapter 5).

### 1.3.2. Electrospray ionisation (ESI)

Its characteristic of "soft" ionisation technique is not the only reason why electrospray ionisation is nowadays the technique that is mostly coupled with FT-ICR MS. Opposite to MALDI, ESI introduces samples in the mass spectrometer as liquids, thus increasing the number of species that can be analysed, as long as they can be dissolved. Also, ESI frequently leads to the formation of multiply-charged ions, which coupled with the high resolving power of FT-ICR MS can further expand the analysis range by reducing the $m/z$ of macromolecules whose mass would normally be over the upper mass limit of the technique when singly charged. As for the MALDI technique, ESI yielded its developers a Nobel prize [29, 30, 31].

The operational arrangement for ESI is shown in Figure 1.10. The liquid sample is exposed to a potential difference that causes charge separation inside the solution. It is then passed through a fine needle with the aid of a syringe pump. As the sample exits the needle, it is exposed to an inert gas, such as nitrogen, which causes the solution to nebulise through formation of a Taylor cone. Droplets form from the Taylor cone, and desolvate until the Coulombic repulsions of the charges inside them overcome their surface tension (the Rayleigh limit). Over the Rayleigh limit, the droplets undergo Coulombic explosions or discharge events with the formation of smaller droplets and eventually ions. A difference of potential is applied between the needle and the instrument front end, of a polarity opposite to the ions of interest (Figure 1.10). The small droplets continue the process of desolvating until only ions in the gas phase are left, and the ions are attracted to the front-end of the instrument and finally introduced into the mass spectrometer [32, 33].

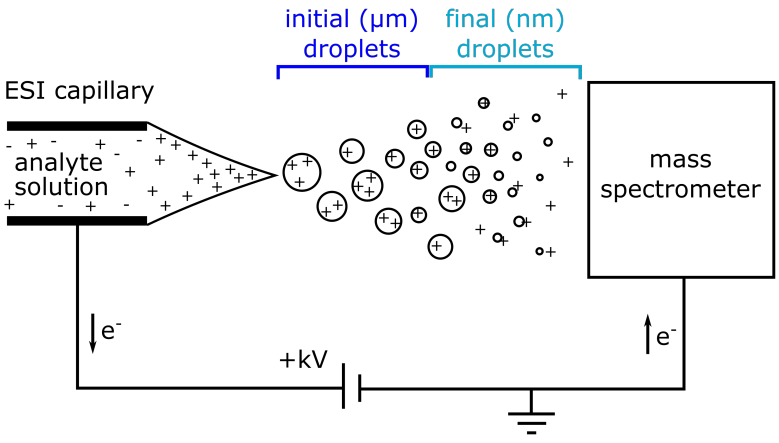



Figure 1.10.: Schematic depiction of an ESI source operated in positive ion mode. A Taylor cone is generated at the tip of the needle, with subsequent formation of droplets, and desolvation until the Rayleigh limit is reached. The generated ions (positive) are attracted by the spray shield of the mass spectrometer inlet, negatively charged.

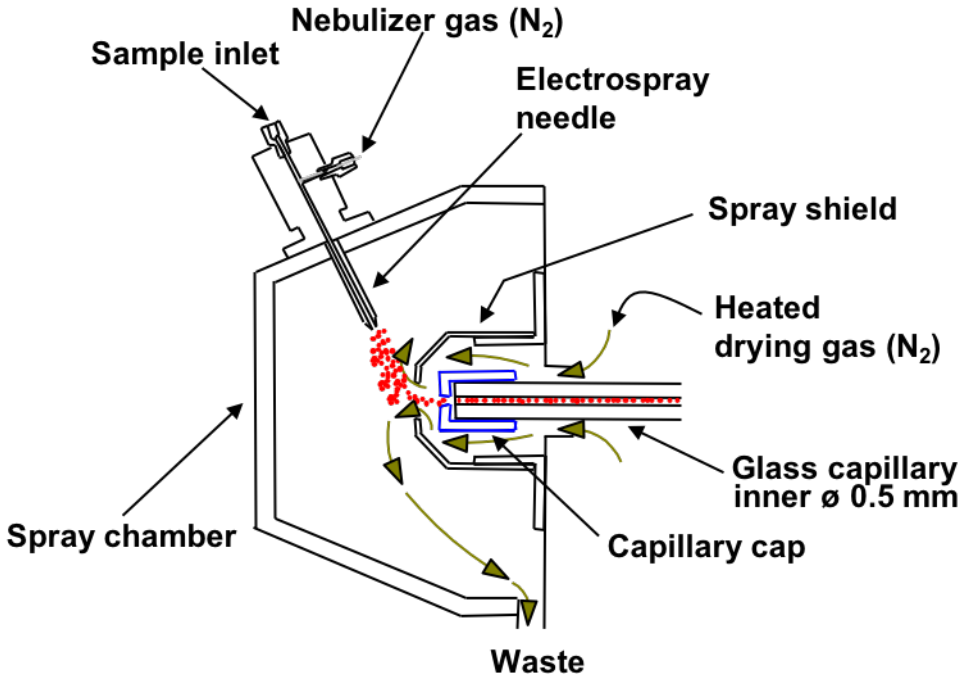



Figure 1.11.: Schematised configuration of ESI for MS experiments in a Bruker ESI source. The sample is introduced through the sample inlet in the electrospray needle, where a potential difference is applied in order to separate charges. When the droplets exit the tip, desolvation is aided by a nebuliser gas (nitrogen, in this case), and the resulting smaller droplets/ions are attracted to the spray shield, charged with an opposite polarity. Finally, the generated ions are introduced into the mass spectrometer. The generated ions are attracted by the spray shield, at an opposite charge. Courtesy of Bruker Daltonics, Bremen, Germany.

Different models are used to explain the desolvation of molecules into the gas phase, the most accredited of which are the ion ejection model (IEM) and the charge residue model (CRM). The models are visually explained in Figure 1.12. The first model suggests that the field outside the droplet eventually becomes sufficient for a small charged analyte to be ejected from the liquid-phase medium (Figure 1.12.a) [34], while the second is based on subsequent evaporation/fission of the droplets into progressively smaller ones, until only one ion on average is left [35, 36, 37]. Analytes enter the gas phase when the droplet evaporates completely, leaving the residue charge on the particles (Figure 1.12.b). Studies demonstrated that small ions form according to the IEM, while large particles (such as folded proteins or protein complexes) are liberated into the gas phase through CRM.

A third mechanism for ion formation in ESI has been proposed by Konermann *et al.*, the chain-ejection model [38]. This model explains the formation of ions of long polymers or unfolded proteins. In fact, due to their elongated structure, such molecules tend to migrate away from the centre of the droplets, by exposing their

most hydrophobic moieties towards its surface. Eventually, the terminus of the polymer/protein is ejected from the droplet, and consequently charged due to its proximity to the droplet surface (where the charges are). This ejection triggers the subsequent release and charging of the rest of the polymeric/amino acid chain, with the ion eventually being liberated into the gas phase (Figure 1.12.c).

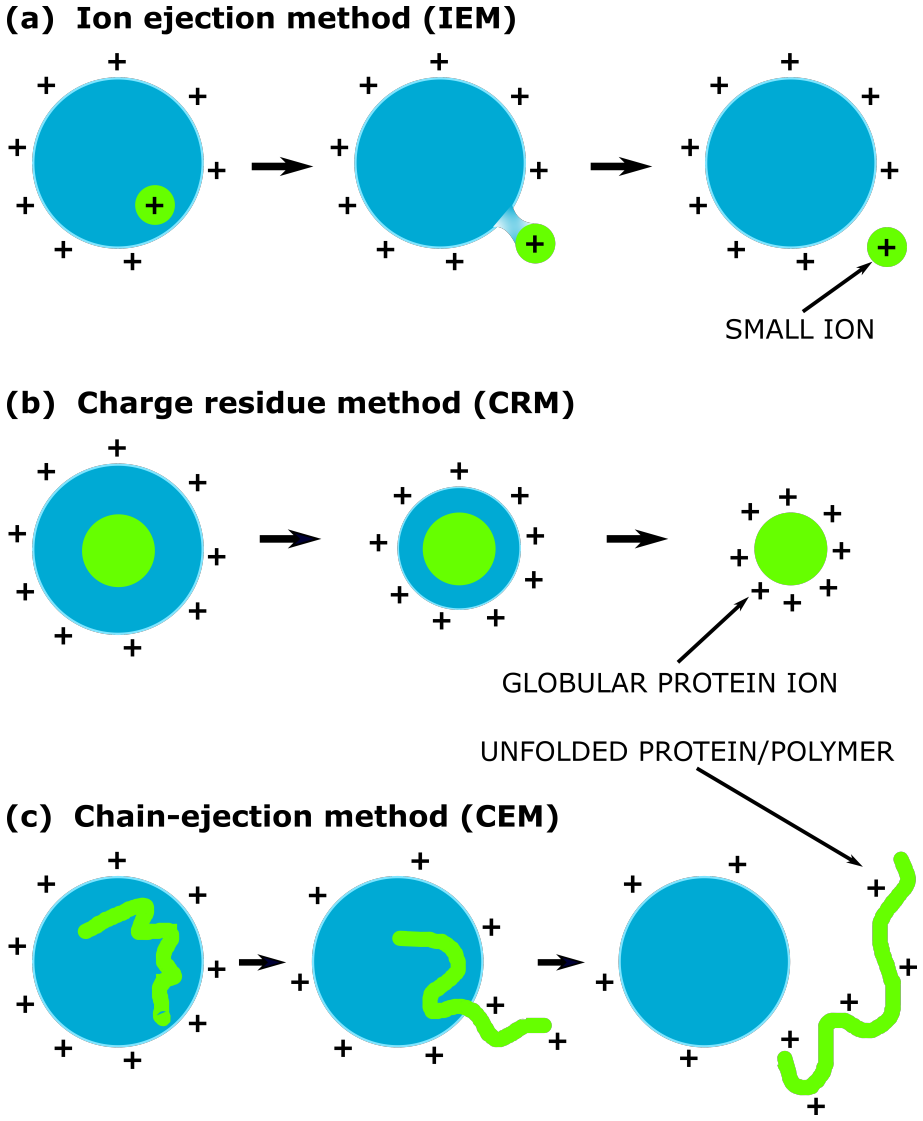



Figure 1.12.: Models used for the evaporation of droplets into the gas phase with generation of charged particles in electrospray ionisation. (a) Ion evaporation model (IEM). (b) Charge residue model (CRM). (c) Chain-ejection model (CEM).

ESI produces both positive and negative ions, and usually the polarity of the applied potential difference is chosen based on the sample, and can be changed by the user to introduce both positive and negative ions alternatively.

Mass spectrometry is based on the identification of species based on their $m/z$. For

higher charges, the mass-to-charge ratio of the molecules under analysis diminishes, allowing - as mentioned - to observe macromolecules whose mass would normally be over the upper mass limit of the technique. FT-ICR MS's resolving power allows to observe the isotopic distribution of the species under analysis, up to their fine structure. A peak is resolved when a clear noise line can be observed between two adjacent peaks. The charge of a species can be determined through the difference in $m/z$ between two peaks of the same isotopic distribution,usually the mono-isotopic peak M and its M+1 isotopologue. The distance between these two peaks corresponds to $\Delta m/z = 1/z$ Daltons per charge, allowing to calculate the charge $z$ of the species. The ability to assign the charge of a species of a certain mass-to-charge ratio allows to assign the molecule, thus allowing for identification of macromolecules of considerable size. The presence of the same molecule at different charge states may complicate the resulting mass spectra, but in case of FT-ICR MS the resulting lower $m/z$ means higher resolution and so the possibility to analyse larger molecules. Most mass spectrometers are able to perform reasonably over $m/z$ 10000, most notably ToF instruments, which can analyse proteins in the ultrahigh mass range ($m/z >$ 20000) at expenses of RP [39, 40], but an optimal experimental $m/z$ range is often obtained at lower mass-to-charge ratios. *E.g.* for FT-ICR MS, the optimal experimental $m/z$ range is between ∼150 and 5000. The ability to add multiple charges to a molecule in order to bring it into the optimal range is a benefit to the entire MS field. One of the main advantages of ESI is also a direct connection with separation techniques such as liquid chromatography (LC), providing hyphenation of the two techniques (LC-MS) for on-line analysis of complex mixtures [41].

The growth of electrospray ionization promoted the growth in FT-ICR mass spectrometry, for FT-ICR mass spectrometers offer the ultra-high resolution which is particularly well suited to coupling with ESI. For these reasons, ESI FT-ICR mass spectrometry has become increasingly important for the study of biological molecules in recent years.

#### 1.3.2.1. nanoelectrospray ionisation (nESI)

In order to provide a lower sample consumption, electrospray ionisation has been scaled down to a much smaller size, with the development of nanoelectrospray ionisation (nESI) [42, 43]. nESI makes use of much smaller ion emitters, reduced to needles with orifices between 1 and 10 µm . As a consequence, smaller droplets are formed, with a major efficiency of the ionisation technique. Due to the reduced droplet size, nESI does not require the aid of a nebuliser gas, and the sample is pushed into the mass spectrometer entirely with the application of a potential difference between the emitter and the front end of the mass spectrometer and the vacuum inside the ion source. Hence, no syringe pump is required. nESI produces much smaller plumes at a very low flow rate (in the range of about tens nL/min)

compared to ESI [44, 45, 46], and the emitters are positioned more closely to the instrument compared to the above-described technique ($>$ 1 mm) [47]. The emitters can be purchased or laboratory-made, usually by pulling borosilicate glass, metal, or silica tips. In any case they are coated with a conductive material, or they are connected to a conducting wire, in order to provide electrical contact with the solution and complete the circuit that attracts the sample out of the emitter. The experimental setup of nESI is shown below (Figure 1.13).

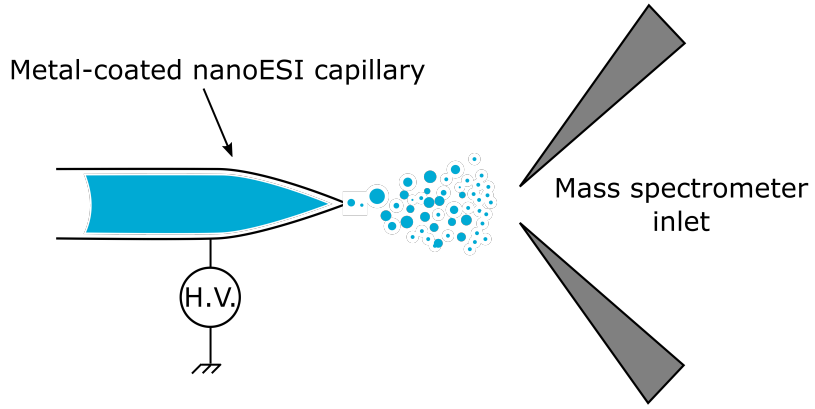



Figure 1.13.: Experimental setting for nanoelectrospray ionisation. The ion emitter is positioned close to the mass spectrometer inlet, due to the smaller plumes generated by the desolvation of the formed droplets. The electrical continuity between the solution and the instrument front end is maintained through a metal coating at the tip of the emitter. .

#### 1.3.2.2. ESI/nESI of proteins

The ability of ESI to generate multiply-charged ions is well suited for the analysis of large proteins, which necessitate of and can accommodate more than one charge in order to be detected in the optimal $m/z$ range. Peptides and proteins contain many functional groups in their amino acids that are able to retain positive or negative charges, most notably the side-chain residues of arginine, lysine, histidine, and the N-terminus of the chain, through addition of a positive adduct (proton or cation), and glutamic and aspartic acid, through deprotonation. In order to be charged, these sites need to be exposed to the solvent, meaning that if the protein is in a particular conformation for which these charging sites are enclosed in its tertiary or quaternary structure, multiple charge states do not occur extensively. Tertiary and quaternary structures refer respectively to the three-dimensional structure of a protein chain or the structure resulting from aggregation of more chains that operate as a sin-

gle functional unit. Protein conformations depend on the environmental conditions, such as solvent, pH, temperature, and so on, and can be manipulated accordingly. The use of mixtures of organic and aqueous solvents, usually water/methanol or water/acetonitrile 50/50 v/v, and the addiction of an acid, usually formic acid or acetic acid, up to 0.3% of the total volume of the solution, helps protein unfolding, making more sites accessible for protonation. The experimental conditions that use solvents and acids in order to unfold ("denature") the proteins under analysis are called denaturing conditions. Usually, a higher charge is preferred in order to analyse proteins with a considerable accurate mass, above all for more informative experiments, such as tandem mass spectrometry (see Section 1.6). In denaturing conditions, ESI-MS of proteins usually exhibits a wide range of charges, reporting in the mass spectrum multiple isotopic distributions of the same protein with different charges and different combinations of adducts. Although this situation might overcomplicate mass spectra with more than one analyte, the presence of proteins in more than one charge state represents an advantage for tandem mass spectrometry, since fragmentation of some proteins depends on their conformation. Therefore fragmenting different charge states provides additional information about their structure [48]. However, the presence of multiple charge states makes that the signal provided by a protein in solution is divided on the spectrum between all its charge states/different adducts, reducing the sensitivity of the technique.

Denaturing conditions are ideal when the primary structure of a protein, *i.e.* the amino acid sequence, needs to be elucidated, *e.g.* for preliminary studies on proteins or *de novo* sequencing. However, the use of harsh solvents and low pH disrupts the most labile bonds, hindering the identification of some post-translational modifications (PTMs) and a more elaborated structural analysis of the tertiary and quaternary structure. Whenever a protein conformation or delicate interactions between protein sites need to be studied, proteins are dissolved in conditions that resemble their native environment, called indeed "native conditions". Native conditions may involve the use of buffer solutions, but not of organic solvents or acids, and ESI temperatures are kept lower compared to the standard ESI experimental conditions. Most buffers used in structural biology are not easily compatible with ESI-MS because the involved salts hinder the protein ionisation process, leading to signal suppression. For its compatibility with ESI-MS, aqueous ammonium acetate is used as a buffer solution for native protein analysis, in concentrations between 5 mM and 1 M [49]. Native conditions are more technically challenging, due to the less effective ESI desolvation and the more folded protein conformations, which limit the availability of the protonation sites [50, 51, 52].

## 1.4. Ion transfer

All sorts of ion manipulations inside an FT-ICR mass spectrometer are performed through the application of electromagnetic fields, from the potential difference applied to attract the ESI-generated ions from the ion source to the mass spectrometer inlet, to the various ion funnels, lenses and guides used to move the ions inside the instrument.

Ion funnels allow to focus the widely spread plumes of ions and neutral gas molecules generated in-source towards a smaller orifice [53]. This phenomenon is achieved through the use of a series of capacitively-coupled metal plates starting with a central wide circular aperture, which becomes smaller with every plate, hence the name funnel. The plates are held at alternative voltages in order to attract or direct ions in the right direction. A more traditional method to achieve the same process as ion funnels is the use of conductive metal cone-shaped funnels called skimmers. Skimmers are still in use in contemporary ion sources for their ability, at high applied voltages, to cause ion-skimmer collisions, which aid the desolvation of large ESI-generated ions and salt-clusters [54]. The variation of the voltages applied to the skimmer surfaces also allows to perform nozzle-skimmer dissociation [55, 56], an in-source dissociation (ISD) which generates fragments similar to other collisionally-activated techniques (See Section 1.6, and Chapter 4 for an application of ISD).

Although ion skimmers and funnels are able to focus and direct ions towards confined parts of the mass spectrometer, they are not able to transfer or store ions in the required UHV conditions for FT-ICR MS. This task is given to multipoles, components made of bars alternatively charged with opposite polarities (Figure **??**). Classical examples of multipoles are the quadrupole, hexapole, and octopole. Although the quadrupole can be used as a mass analyser [57], its ability to discriminate between masses is here used to select specific masses, allowing to manipulate the ion population of the ICR-cell. Quadrupoles are used to select small ranges of $m/z$ by stabilisation of the selected masses and destabilisation of the others, which would hit the rods, lose charge, and be pumped out of the instrument as neutral gas molecules (Figure 1.15) The selected ions can subsequently be stored and accumulated in another multipole, such as an hexapole. Principles of the stabilisation of the ion transfer through multipoles are given by the Mathieu stability diagram (Figure 1.15.a). By varying the values of $a$ and $q$, ions are brought in or out the stability area. Ions in the stability area are successfully transmitted as per trajectory similar to Figure 1.15.b. Ions outside of the stability area are not transmitted, as per Figure 1.15.c. The selection of a certain range of mass-to-charge ratios of interest is called isolation (in this case, "quadrupole" or "quadrupolar isolation"). Ion accumulation in a trap contributes to manipulate the total number of ions that enter the ICR-cell, and it is a fundamental part of tuning. The possibility to isolate ions before they enter the

ICR-cell is also a great advantage for tandem mass spectrometry, because it allows to fragment one species at a time. Hexapoles and octopoles are more effective ion guides than quadrupoles, and the former can be used as collision cells for the ions selected through quadrupolar isolation.

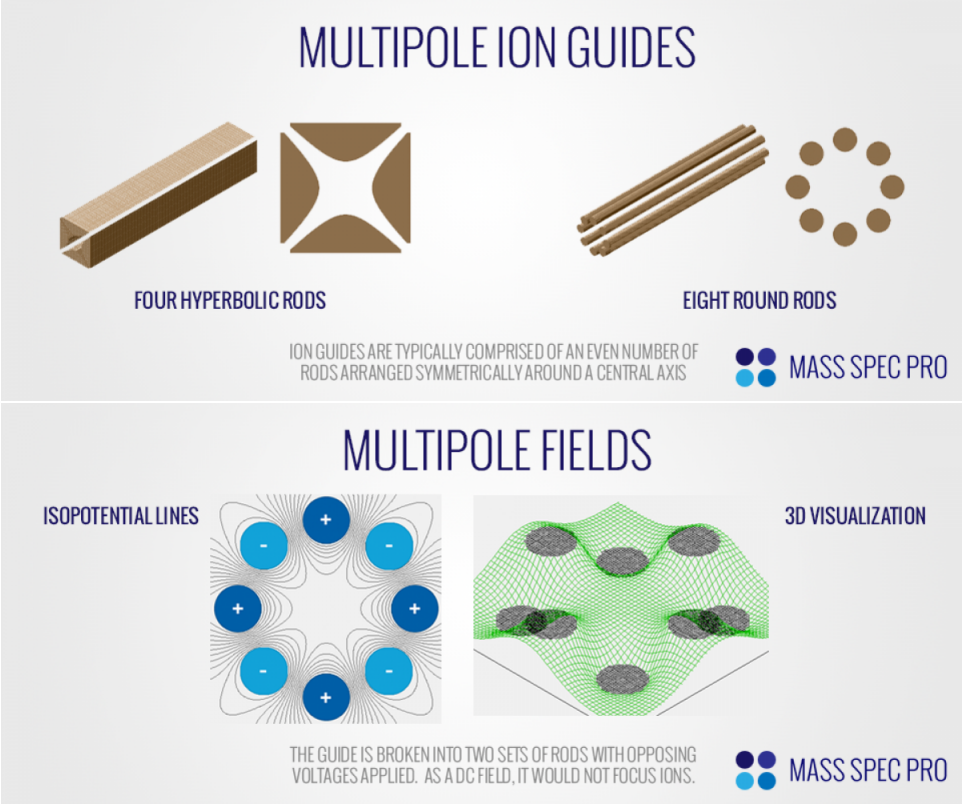
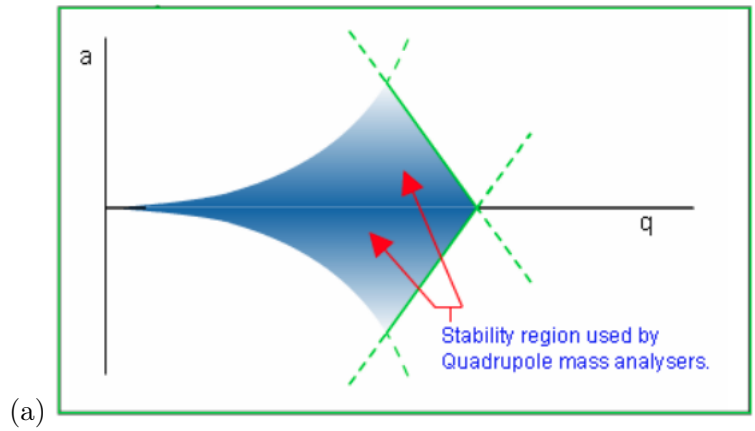
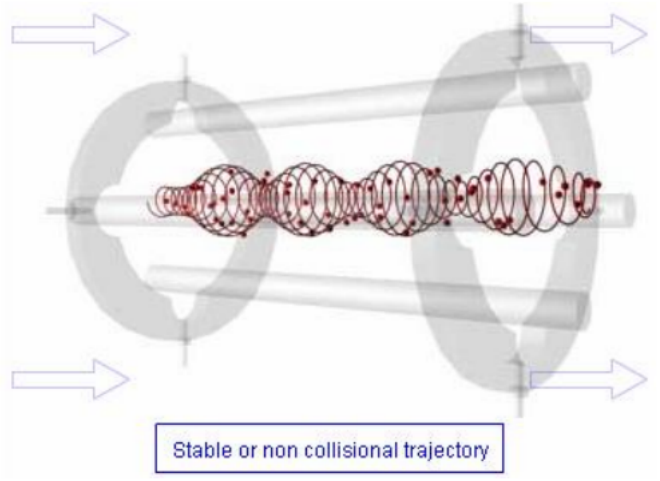


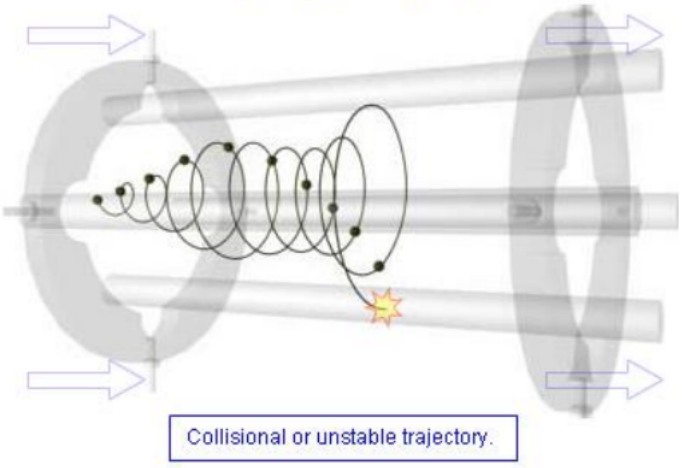

Figure 1.14.: Examples of mutipole ion guides and relative ion fields. Top: quadrupole composed of four hyperbolic rods, and octopole with cylindrical rods. Bottom: electric field inside an octopole, showing the trapping potential in a perpendicular section and as a tridimensional visualisation. Ions are confined in the central part of the ion guide, oscillating around the "well" of lowest potential. NB. The rods alternate polarity as a function of time. Reproduced from http://www.massspecpro.com/technology/ion-optics/multipole-ion-guide, accessed 16/08/2017.




(a)




(b)




(c)

Figure 1.15.: Ion trajectories inside a quadrupole. (a) Stabilised ions will travel the entire quadrupole range. (b) Destabilised ions will eventually expand their trajectory until they hit the quadrupole rods, with consequent loss of charge, and impossibility to be manipulated towards the mass analyser. Reproduced from chromacademy.com (http://www.chromacademy.com/lms/sco36/Fundamental_LC-MS_Mass_Analysers.pdf), accessed 17/02/2018.

## 1.5. ICR-cells

### 1.5.1. Cell geometry and development

ICR-cells of cubic geometry were the first cells to be built and used for many years on successful experiments [58, 59], despite their main drawbacks of low ion capacity and not being optimised for superconducting magnet geometries. Many ICR-cells designs followed the cubical cells with time, each act to overcome different challenges of the FT-ICR MS experiments. The main developments in geometry of ICR-cells are shown in Figure 1.16 after the cubic design (Figure 1.16.a)

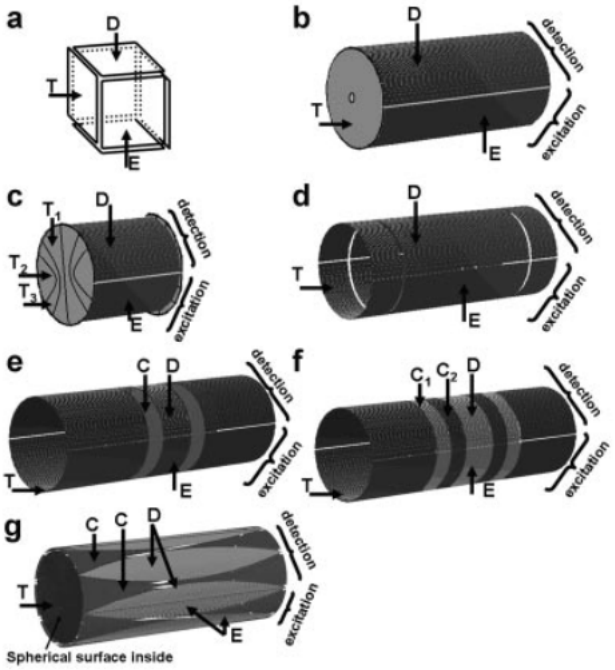



Figure 1.16.: ICR-cell geometries. a) Cubic cell. b) Cylindrical cell. c) Infinity cell. d) Open cylindrical cell. e,f) Compensated cylindrical cells with different numbers of "shimming" electrodes. g) Dynamically harmonised cell. D = detection. E = excitation. T = trapping. C = compensation. Reproduced from Nikolaev *et al.* [60].

The closed cylindrical cell (Figure 1.16.b), with two pairs of excitation and detection electrodes, was developed to optimise the magnetic field linearity inside the cell, when enclosed in the superconducting magnet bore. The curved shape of the detection electrodes allowed their more prolonged exposition to the ions orbiting inside the cell, improving sensitivity. The closed cylindrical cell, although optimised in shape for the magnet hardware and better detection, suffered from the presence of the trapping plates, which caused electric field permeation through the cell with subsequent perturbation of the ion motions at its center. Open-cell geometries such as the cylin-

drical cell of Figure 1.16.d were developed to resolve this problem, but suffered from ion ejection towards the z-axis ($z$-ejection) following the excitation event [61]. $z$-ejection was subsequently resolved by capacitively coupling the excitation electrodes to the trapping electrodes [62]. Figures 1.16.e and f represent compensated open cylindrical cells based on the models proposed respectively by Tolmachev *et al.* [63] and Bruskern *et al.* [64]. The purpose of the original compensated cylindrical cell was to minimize the effects of the nonlinear components of the electric field inside the cell. The theory of FT-ICR MS implies the motion of ions in a theoretically infinitely linear magnetic field, and non-linearities introduce variations in the cyclotron frequencies based on the position of the ion in the cell, leading to a reduced resolving power and mass accuracy . The cell is a cylindrical vessel constituted by four identical sides which contain, starting from the base of the cylinder, two external end-caps, two inner trapping plates, two sets of compensation (or shim) electrodes, and the excitation or detection plate. In the capacitively-coupled open cylindrical cell, excitation and detection electrodes are capacitively coupled in order to reduce the $z$-ejection during the excitation event [65, 66].

In order to avoid the problems caused by end-cap electrodes and a cell of finite dimensions, Caravatti *et al.* introduced the Infinity cell (Figure 1.16.c). The Infinity cell is a closed cylindrical cell with special trapping electrodes, made of wires mounted on a ceramic circular surface with a design that mimics the behaviour of an infinitely long ion trap. The cell virtually eliminates the $z$-ejection, with consequently improved sensitivity compared to the open cylindrical cell. The main disadvantages of the Infinity cell are the required extensive tuning, the small ion capacity, and a high dependence of the performance from the cell alignment, due its small size and close geometry, which cause challenges when laser/electron beams are used for extended experiments. Nonetheless, the Infinity cell has been for years the main core of the Bruker Daltonics FT-ICR mass spectrometers, achieving RP $> 10^6$ and MMA close to ppb [12].

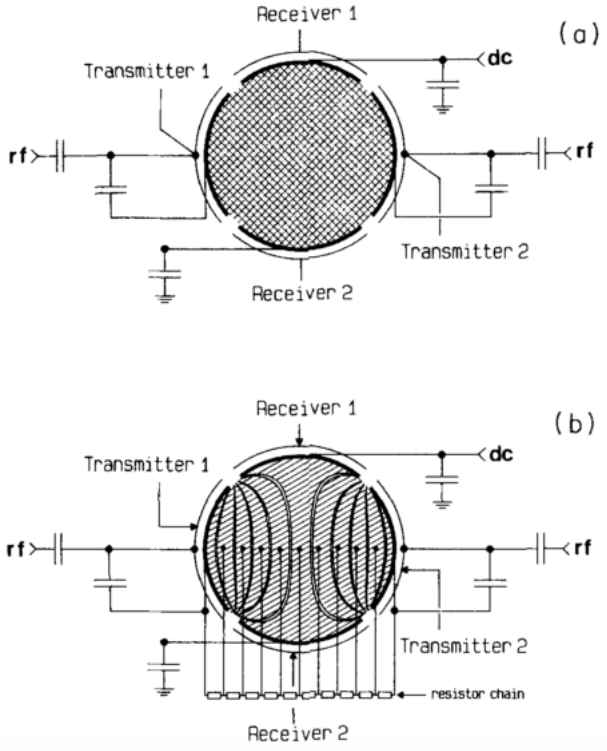



Figure 1.17.: Scheme of a transversal section of the Infinity cell. Top: inside of the Infinity cell, with associated circuitry. The electrodes named "Transmitter" are the excitation electrodes, while the "Receiver"s are the detection electrodes. Bottom: trapping plate of the Infinity cell. The wires mounted on a ceramic surface mimic the behaviour of an infinitely long ion trap. Reproduced from Caravatti *et al.* [18].

Figure 1.16.g shows the dynamically harmonised cell by Nikolaev *et al.* [67, 68, 69, 70, 60]. The dynamically harmonised cell, or para-cell, is based on the concept of electric potential space-averaging via charged particle cyclotron motion, permitting an electric force distribution along its cylindrical shape to be equal to the distribution that would be created in an ideal ion trap. The main particularity of this cell are the unusual "leaf-shaped" electrodes. The para-cell reached a RP of over 20M on ions of reserpine ($m/z$ 609) and more than 1M on highly charged BSA ions ($m/z$ 1357) already in its initial sperimentation in 2011 [68].

Development of ICR-cells continues to be a major focus of FT-ICR MS research, with the production of more and more particular cells with singular design goals, trapping electric field shapes, and electrode designs. An example of which is the narrow aperture detection electrodes (NADEL) ICR-cell, which makes use of segmented trapping rings, grid excitation electrodes, and narrow detection electrodes positioned radially inward of the cell's axis (Figure 1.18) [71]. The NADEL cell offers an improved electric field distribution due to the decrease in disturbance created by the insertion of the narrow detection electrodes, and, among other advantages, the

possibility to record simultaneously more than one transient when more NADELs are paired inside the cell. Such studies had been performed before only by coupling more ICR-cells in series.

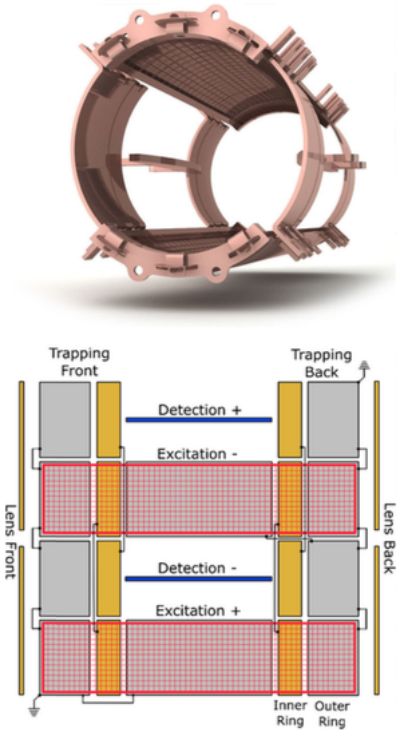



Figure 1.18.: The NADEL ICR-cell. Top: tridimensional structure of the NADEL ICR-cell, showing the grid-shaped excitation electrodes on the top and bottom of the cylinder, and the narrow aperture detection electrodes, inserted transversally into the cell body. Bottom: scheme of the NADEL ICR-cell, showing the design of the four quarters of the cylinder. Reproduced from Nagornov *et al.* [71].

Finally, in an attempt to improve the availability of ICR-cells to a wider range of users, the Pasa-Tolic group developed such mass analysers using only printed circuit boards (PCB) [72]. The approach was successful, but the PCB-made ICR-cells showed a lower tolerance to UHV compared to the standard ones. The use of PCBs to build ICR cells sparked the idea to build cheaper ICR-cells to continue the multiple-detection experiments mentioned above [73].

### 1.5.2. Ion excitation/detection in the ICR-cell

The excitation/detection procedure performed in the ICR-cell has been briefly explained in Section 1.1.1 and its here reported in Figure 1.19 for a schematised cylindrical cell.

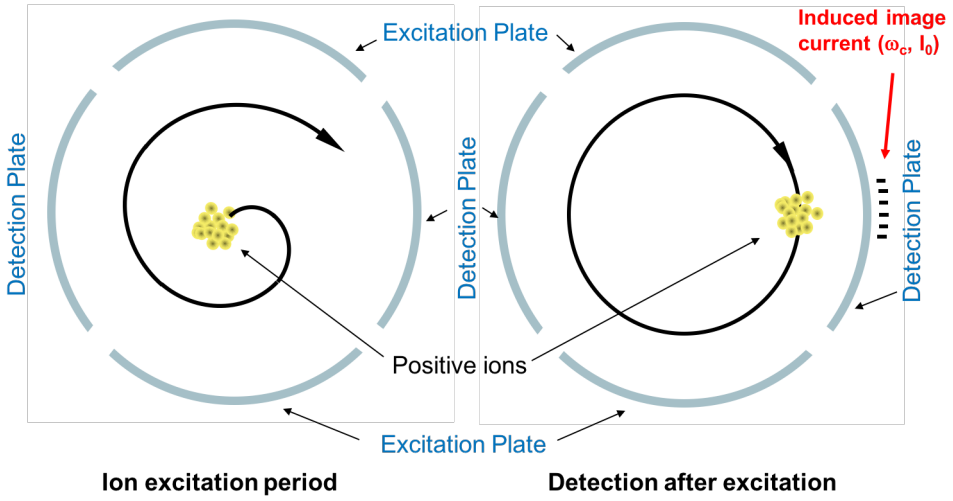



Figure 1.19.: Excitation/detection shown for a cylindrical cell. Left: The ions precess along the magnetic field axis according to their cyclotron frequencies, when a radio-frequency pulse emitted by the excitation electrodes transfer energy to their motion. The ions are excited to higher cyclotron radii independently from their kinetic energy. Right: the coherent ion packets transit close to the detection electrodes, causing the charges of the opposite polarity to move across the plate, generating an image current. After a phase of π, the ion packet will cross the opposite detection electrode, generating the same current, and repeating this process until the ion motions are less stable. The frequency with which the ions approach the pair of detection electrodes is equal to the (reduced) cyclotron frequency. The image current is subsequently recorded, giving the raw data in the domain of time. Courtesy of Bruker Daltonics, Bremen, Germany.

Although excitation of ions to a higher cyclotron orbit produces coherent ion packets, the ion motion within the cell is not ideal. Despite studies show ion confinement within an ICR cell for hours (and days), ions excited to a larger than natural orbit will return to the centre of the cell or will be knocked out of the stable area of the trap usually within several seconds, or minutes [74]. Ion instability can cause the signal to "crash" during the acquisition, badly affecting the resulting mass spectra. This adds to the reasons why the above-mentioned use of UHV and a fine tuning by the user are necessary to achieve the highest performance possible [70, 75].

#### 1.5.2.1. Broadband and Narrowband mode

Although excitation of a single ion packet might seem a simple concept, excitation of a wide range of frequencies, hence a wide $m/z$ range, is not so trivial, and it is achieved via a frequency "chirp". The chirp amplitude raises from lower frequencies and is kept almost constant until it reaches the highest frequencies (low $m/z$), which constitute the low mass cut-off of the excitation. This is called "broadband chirp". Figure 1.20

shows the different profiles of the pulses used to excite different ions, as a function of time (left) and frequency (right). The broadband chirp is reported in Figure 1.20.c, where it can be seen that the frequency profile is not exactly flat, possessing two characteristic "horns" at the beginning and the end of each pulse. Although it might seem that broadband excitation would be limited only by electronics, these anomalies in the frequency pulse produce irregular behaviour and peak shapes in the corresponding $m/z$ regions of the resulting mass spectrum [5, 6].

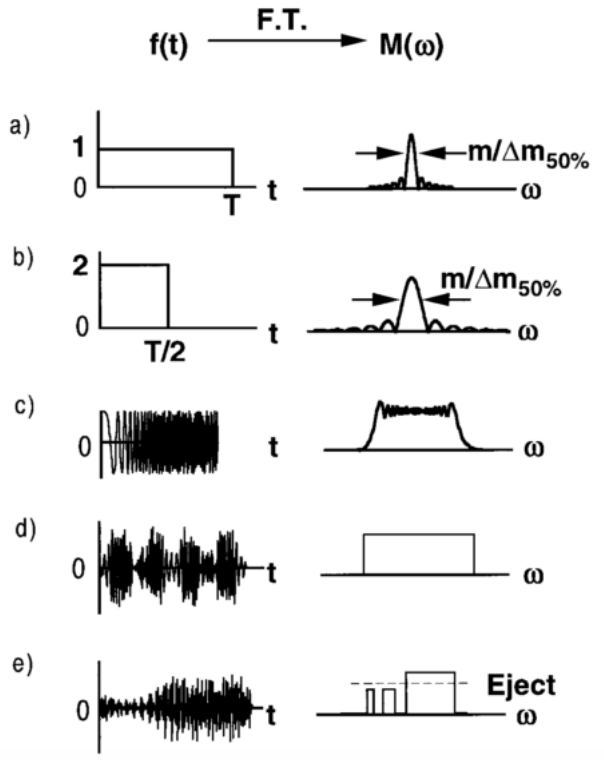



Figure 1.20.: Excitation waveforms in the domain of time (left) and frequency (right). a-b) Rectangular excitation waveforms. c) "chirp" excitation waveform. d-e) SWIFT waveforms, respectively with a continuous wide range of frequencies, and with a notch profile. The second is used as in-cell isolation method. Reproduced from Marshall *et al.* [6].

Narrowband or heterodyne mode offers very high resolution in a much less broad range of $m/z$, by using more data points per frequency. As a result, much longer transients are recorded, with a considerable increase in resolving power for the resulting mass spectrum. Heterodyne mode is based on the use of a reference frequency alongside the recorded transient. The reference frequency is chosen on the base of the short $m/z$ range to analyse, and it is then summed to the transient. The composite signal is then passed through a low-pass filter, which produces a signal given by the sum and difference of the main and reference signals. This heterodyne signal can be analysed at a much lower rate compared to the correspondent broadband signal, resulting in the mentioned longer transients, directly proportional to RP. Narrowband

mode is currently less used than broadband mode, but continues to be an effective way to produce ultra-high resolution mass spectra [76].

#### 1.5.2.2. Stored waveform inverse Fourier transform (SWIFT)

Stored waveform inverse Fourier transform (SWITF) is an alternative way to excite ions in the ICR-cell, developed by Marshall *et al.* [77]. It predicts the frequency profile required for a rectangular excitation, and calculates the inverse Fourier transform in order to obtain the time domain signal required to achieve the excitation profile needed. The excitation profile of SWIFT in the frequency domain (Figure 1.20.d) is flatter, and thus more accurate than the chirp profile. SWIFT has been successfully used to excite ions for detection [5], and also to isolate them inside the ICR-cell (see below). SWIFT can also excite multiple chosen frequencies using a selective excitation profile, allowing to monitor chosen ions of interest without exciting the entire ion population inside the ICR-cell.

### 1.5.3. In-cell isolation (ICI)

The ability of the ICR-cell to selectively excite ions to higher cyclotron radii allows the possibility to selectively eject these ions by exciting them to a radius equal or higher than the cell radius. The ions will then hit the electrodes and be neutralised, with the consequence of being pumped out of the cell as neutral gas molecules. As a consequence, it is possible to excite unwanted ions to force their ejection from the ICR-cell, while "isolating" ions of interest by exciting them to a lower radius for detection. The procedure for which a population of ions is selected for analysis while the rest is ejected from the ICR-cell is called in-cell isolation (ICI).

#### 1.5.3.1. Resonant ejections

ICIs can occur through a broadband pulse of every undesired species with a small window centered on the ion of interest, or through the application of a small, defined, pulse centered to a single or multiple unwanted ions. These procedures are known respectively as "sweep" and "shot" excitations/ejections. Multiple shot excitations in a single analysis require additional time, since each "shot" is applied individually. Since the resonant ejections use a chirp excitation waveform, they have the main disadvantages of a non-ideal excitation at the edges of the pulse. This might result in isolating/ejecting the species of interest, but it might also eject the species at the edges of the frequency/$m/z$ range.

#### 1.5.3.2. SWIFT for ICI

The ability of SWIFT to excite the ions in the ICR-cell similarly to chirp pulses allows the excitation technique to be used as ICI. It is in fact possible to calculate an ejection profile in order to create a pulse which ejects the unwanted species isolating the ions of interest in the ICR-cell (Figure 1.20.e). The results are the same as applying sweep and/or shots, with much more accuracy, due to the more ideal frequency excitation profile of SWIFT compared to chirp. SWIFT has been shown to produce high-resolution in-cell isolations [78, 79].

#### 1.5.3.3. Correlated harmonic excitation fields (CHEF)

By using resonant ejections of unwanted ions, it is possible to transmit some excitation energy to the ions of interest. This off-resonance excitation can cause unwanted ejection of the ions of interest, reducing the resolution of the in-cell isolation and the sensitivity of the technique. Correlated harmonic excitation fields (CHEF) is a method that allows very efficient in-cell isolation by on-resonance excitation of unwanted species and simultaneous off-resonance of the ions of interest [80]. In this way, the technique prevents the target ions to gain energy and incur in eventual ion loss. CHEF calculates the excitation of the unwanted ions and adjusts the excitation time $T_{exc}$ or field magnitude $V_{pp}$ to de-excite the ions of interest the same number of times that they would be off-resonantly excited. So, while the unwanted ions see their radii increased up to ejection, the ions of interest are kept at the centre of the cell. The user can set one of the two main parameters, and the other can be varied. Amongst the main advantages of CHEF, there are its application to many species at a time by modifying the waveform used, with the addiction of each target (multi-CHEF), the extraordinary isolation resolving power, and the virtually zero-ion loss during isolations. The main disadvantages of the technique rely upon the requirement of long CHEF pulses, with the consequent increase of the duty cycle and eventual ion loss for high performance FT-ICR MS instruments, even if nowadays it does not constitute a significant problem; and there is the hypothesis that it slightly alters the position of isolated species after ion-activation in tandem mass spectrometry.

## 1.6. Tandem mass spectrometry

Accurate mass measurement is not enough to assign a molecular structure, since mass alone does not define an unknown molecule nor can discriminate between structural isomers. A given elemental composition can virtually have an infinite number of structures, and an accurate mass measurement can correspond to different atomic combinations. To address this issue, it is possible to obtain more extensive structural

information from a molecule by fragmentation of such molecule and assignment of diagnostic fragment ions. This procedure is called tandem mass spectrometry, or MS/MS. In tandem mass spectrometry, a species is usually separated from a mixture before introduction into the mass spectrometer, or isolated from the mixture through quadrupolar isolation, and fragmented individually. The result is a tandem mass spectrum, *i.e.* a spectrum of all the fragments produced by the formerly isolated ion, called precursor ion. The procedure is shown in picture 1.21.

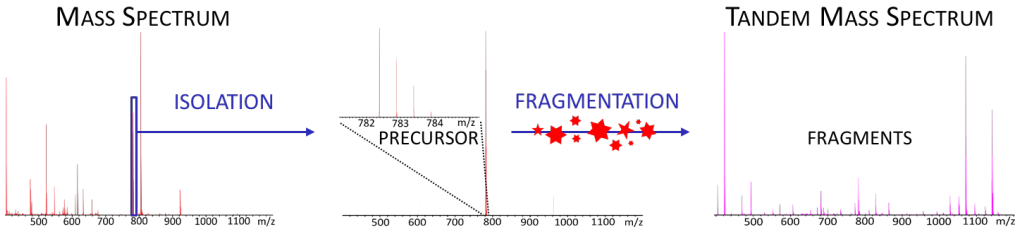



Figure 1.21.: Workflow of tandem mass spectrometry illustrated on a mass spectrum of tryptic digested calmodulin. A substance of interest is isolated from the mixture on the right spectrum, and subsequently fragmented, producing a mass spectrum of all the fragments generated by the isolated precursor. Left: mass spectrum of tryptic digested calmodulin. A peptide of interest is selected for isolation. Centre: isolation spectrum of the precursor of interest. The insert shows the high resolution of the obtained quadrupolar isolation, containing only the isotopic distribution of the selected peptide. Right: tandem mass spectrum. All the peaks in this spectrum correspond to fragments of the isolated precursor, plus the precursor peak. In other words, a tandem mass spectrum shows the fragmentation pattern of the activated precursor.

Furthermore, in order to retrieve even more information from an analyte, a fragment ion can be further isolated and fragmented up to $n$ times, in a tandem mass spectrometry procedure called $MS^n$.

Tandem mass spectrometry is one of the most powerful methods for the acquisition of structural information about a molecule, and it has shown to be a versatile method for the analysis of different molecules and macromolecules, amongst which peptides and proteins, polymers, and sugars, which are the main analytes treated in the work of this Ph.D. thesis.

Fragmentation/dissociation occurs via molecule energization or reaction. Energy can be added to precursor ions in a wide range of ways, *e.g.* in small steps primarily in vibrational modes, as in the case of low energy collisional activation, or in a single event, such as the absorption of UV photons. The rate and amount of energy deposition, as well as the different mechanics involved in the interaction between the dissociation technique and the analytes, affect the outcome of tandem mass spectrometry, in terms of fragmentation efficiency, and types and abundances of the

ions produced. Several methods exist to produce fragment ions by adding energy to their molecular structures, including methods involving collisions with inert gases or surfaces, interactions with electrons or electron-donating/-accepting reagents, or absorption of photons. Many of these activation methods generate different kinds of ions, and are complementary to one another, and the use of more than one activation method to a certain analyte can provide more supporting information to structure assignment. At present, no activation method has proven to be universal for all the mass spectrometry platforms or compound classes, even if FT-ICR mass spectrometers retain the highest versatility regarding MS/MS. Exploration of existing activation methods and development of novel activation techniques are a research focus for the entire mass spectrometry field.

## 1.7. Approaches of mass spectrometry to protein and peptide analysis: top-down and bottom-up proteomics

Tandem mass spectrometry is a key feature of mass spectrometry for the study of structures, functions, and interactions of proteins and peptides at different physiological levels (proteomes), called proteomics. Mass spectrometry offers two main approaches to proteomics, called respectively bottom-up (BUP) and top-down (TDP) proteomics.

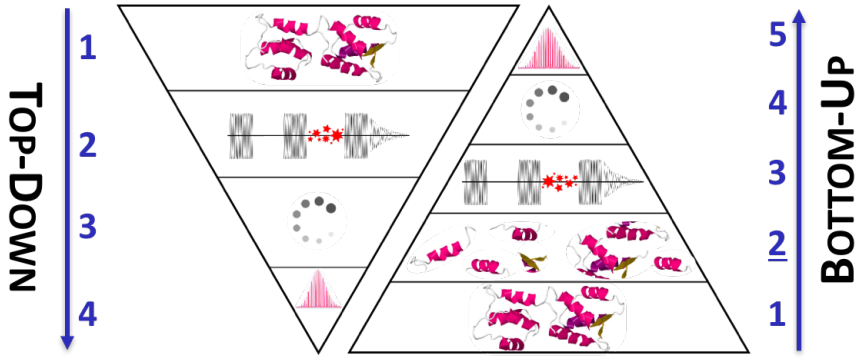



Figure 1.22.: Top-down and bottom-up approaches. Left: in the TDP approach, proteins are dissolved (1) and fragmented (2). Data are then processed (3) and a top-down tandem mass spectrum is generated (4). Right: the intact protein (1) is enzymatically digested into smaller peptides (2), before fragmentation (3), data processing (4), and generation of a bottom-up tandem mass spectrum (5).

The bottom-up approach is a popular method for which a protein is enzymatically digested into smaller peptides, which are representative of its original structure. Depending on the protein size, the generated peptides can be separated through liquid

chromatography (LC) or directly infused into the ionisation source of the mass spectrometer, for subsequent ionisation, eventual fragmentation, and detection. Sometimes, proteins can be assigned based on the accurate assignment of known representative peptides (peptide fingerprinting). Otherwise, peptides can be isolated, activated, and structurally elucidated through the analysis of their fragmentation pattern. BUP is generally easier compared to TDP because peptides are generally easier to fragment compared to entire proteins, and generate much less complex spectra. Furthermore, it is not always necessary to fragment a high number of peptides in order to assign a protein structure. On the other hand, the harsh conditions involved in the enzymatic digestion disrupt all the labile residue interactions and most of the post-translational modifications, with complete loss of the tertiary and quaternary structure information [81]. Furthermore, detection of the peptide ions depends on their ionisation efficiency, and not all the peptides in a mixture ionise in the same way, with the result of the absence of some ions, very high mass-dynamic ranges (some peptides with very low abundances while others at high abundances), and peptides present with more charges and adducts. To quote Brian T. Chait, "It is a little like having a jigsaw puzzle, where many of the pieces are missing" [82].

The top-down approach, on the other hand, involves the analysis of proteins starting from their entirety. Proteins are dissolved in denaturing or native conditions, and fragmented through tandem mass spectrometry. Their structure is then elucidated through the analysis of the mass-to-charge ratios of the generated diagnostic fragments. TDP retains the tertiary/quaternary structure information of the protein under analysis, such as PTMs, but it presents more technical challenges. Protein-level separations are harder to perform, and top-down mass spectra are generally at a high level of complexity. Nonetheless, for the amount of information generated, TDP tends to be preferred and it is now at the centre of the development for proteomics analysis.

Another, less popular mass spectrometry approach to proteomics is known as middle-down proteomics, and involves the enzymatic digestion of molecules into peptides that are larger than the BUP peptides but smaller than the intact proteins [83, 84]. This method was developed because of the technical challenges imposed by TDP .

Protein assignments are performed by analysis of the generated fragmentation patterns, with the use of dedicated databases, or algorithms for *de novo* sequencing. The former method is applied to look for or assign known proteins, while the latter is designated to the identification of unknown macromolecular sequences. A protein is assigned or identified when a certain percentage of its structure is elucidated. Amino acid sequences are found through the analysis of diagnostic ions, and results are given with the percentage of known inter-residue cleavages: for known proteins, this percentage is called *cleavage coverage*. When a peptide/protein is identified from scratch (*de novo*), and its structure is built through the number of identified

ions/cleavages found, such percentage is called *sequence coverage*.

### 1.7.1. Fragmentation nomenclature for biomolecules

Systematic nomenclature has been developed to categorise the formation of different fragment ions, with the use of alphabet letters, and numbers as subscript/superscript.

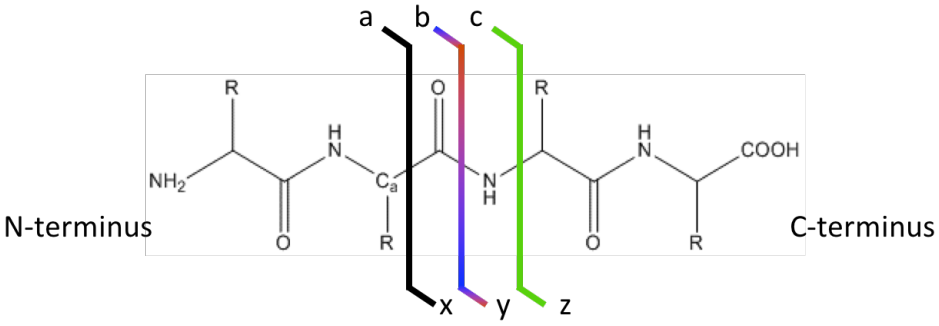



Figure 1.23.: Fragmentation nomenclature for proteins and peptides, as developed by Roepstorff *et al.* [85].

For proteins and peptides, the fragmentation nomenclature currently in use has been developed by Roepstorff *et al.*, and it is illustrated in Figure 1.23. The figure shows the three main diagnostic bond cleavages that can occur on a peptide following different fragmentation techniques. *E.g.* low energy collisional activation techniques such as collisionally-activated dissociation (CAD) or photodissociations such as infrared multiphoton dissociation (IRMPD) have the ability to break the amide bond between peptides, *i.e.* the C-N bond. The fragmentation mechanism leads to the formation of two types of ions, $b$ and $y$, if there is enough charge to detect the ion pair. Depending on the available charge on the peptide, and on the ionisation efficiency of the generated fragment, fragment ion abundances can vary, bringing sometimes to the identification of $b$ or $y$ ions only, if the charge is retained on the N- or the C-terminus respectively. Small neutral losses (such as water and ammonia) and the generation of $a$-ions (by loss of $CO_2$ from $b$-ions) constitute a minor path of low energy activation techniques. On the other hand, electron-based techniques such as electron-capture dissociation (ECD) and electron transfer dissociation (ETD) generate $c/z$ ions by breaking the $C_\alpha$-N bond of proteins and peptides, and higher energy activations such as ultraviolet photodissociation (UVPD) lead to $a/x$ ions, other than the other two types.

The nomenclature is based on the broken backbone bond and the terminus that retains the charge, with alphabet letters indicating the cleavage of the $C_\alpha$-C, C-N, or $C_\alpha$-N ($a$, $b$, $c$ for the N-terminus, and $x$, $y$, $z$, for the C-terminus) and numerical subscripts indicating the number of amino acids in the peptide fragment. *E.g.*, the

ion $y_3$ indicates a fragment ion generated from a protein or peptide by cleavage of the amide bond after the third amino acid from the C-terminus.

Extensive fragmentation can lead to the cleavage of more than one backbone bond in a protein/peptide, generating internal ions. Internal ions lose both the characteristic C- and N-terminal portions, and are of difficult interpretation. They can also be generated willingly by isolating a fragment ion and further fragmenting it in $MS^n$ experiments. Other less diagnostic ions include immonium ions, low-mass ions of single amino acid residues, and ions that undergo side-chain losses.

A similar systematic nomenclature has been developed by Domon and Costello for the characterisation of fragments generated with tandem mass spectrometry of glycans [86]. The nomenclature is shown in Figure 1.24 and it is based on the broken bonds and the terminal portions that retain the charge, in this case the non-reducing end, conventionally drawn to the further left, and associated with letters from the initial part of the alphabet, $A$, $B$, and $C$, and the reducing end of the glycan, right, associated with letters from the final part of the alphabet $X$, $Y$, $Z$. The numerical subscripts indicate the number of monosaccharides included in the fragment, starting from the terminal portion. The superscripts on the left part of the letter indicate cleavages inside the carbohydrate ring.

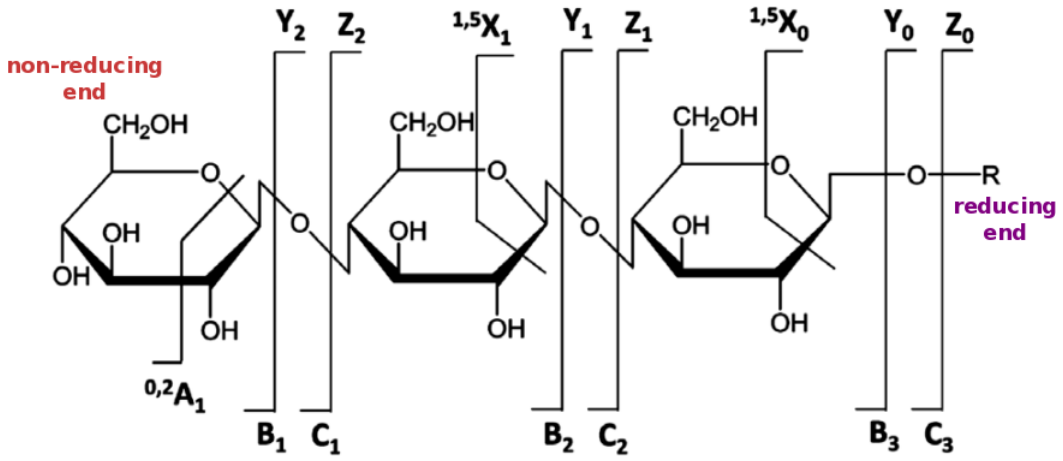



Figure 1.24.: Fragmentation nomenclature for glycans. Adapted from Domon and Costello [86].

### 1.7.2. Collisional activation

Collisionally-activated/induced dissociation (CAD or CID) involves the use of collisions between the analyte ions and an inert gas, such as nitrogen, argon, or helium. In the process, some kinetic energy of the ion is converted to internal energy, which is accumulated until the dissociation occurs [87, 88]. This stepwise, multi-collision process tends to limit the deposition of energy on the ion population, leading preferably

to lower energy dissociation pathways. As a result, many structurally uninformative losses, such as water, ammonia, and carbon dioxide, occur for peptides and proteins, consequently overcomplicating the resulting mass spectra [89, 90]. Furthermore, labile bonds often include post-translational modifications such as phosphorylation [89], which are lost as neutral molecules, and their identification is precluded to CAD. Nonetheless, CAD's high fragmentation efficiency represents an enormous advantage, making it the favourite choice for many applications involving different mass spectrometers.

CAD is a robust technique, well-supported by intensive mechanistic studies [89, 90]. In particular, the *mobile proton* model provided a base to understand the major fragmentation pathways occurring in collisionally-activated dissociation of positively-charged peptides. The model is based on the realistic assumption that the ionising protons are located at the most basic sites of the amino acid chain, such as the N-terminus, or the side-chains of arginine, histidine, and lysine (see Section 1.3.2.1). Once CAD starts, the mobile protons can move to less basic sites, in particular the amide bond, initiating the fragmentation mechanism and leading to the C-N break that precedes the formation of b/$y$ ions. According to this model, non-specific cleavages can occur when the macromolecule is charged with more protons than the number of basic sites it possesses [91, 92].

Ions at high charge states exhibit preferential cleavages of the amide bond N-terminal to a residue of proline, while cleavage of the amide bond C-terminal to an acid residue is more common for ions with less mobile protons. These preferential cleavages might restrict the amount of information provided by the technique for peptides or proteins containing many atoms of proline or aspartic acid, respectively.

### 1.7.3. Electron-based dissociation

Electron-based dissociations (ExD) include all the activation methods involving the interaction with electrons, or electron-donating or electron-accepting ion species. In particular, electron-capture dissociation (ECD) [93, 94, 95] and electron-transfer dissociation (ETD) [96] are the most popular ExDs, due to their extensive applications in the field of proteomics.

ECD is based on the interaction of multiply-charged positive ions with low energy electrons, a process that leads to the formation of a charge-reduced species (CRS), and subsequent fragmentation [97, 98, 99]. It is usually associated with FT-ICR MS, where cations and electrons can be trapped simultaneously in a way challenging to other ion traps, due to the impossibility to trap electrons within stable confinements obtained with radio-frequency voltages. ETD was developed for the implementation of an ECD-like technique in other mass spectrometers, and it involves the reaction of multiply-charged ions with radical anions, causing electron transfer [100, 101, 102].

For both the above-mentioned ExDs, it is thought that the fragmentation mechanism is initiated by capture/transfer of an electron by a mobile proton, with the generation of a hydrogen radical.

The mechanism of ECD has been broadly studied [103, 104, 105]. The CRS formed by capture of a low energy electron undergoes vibrational energy redistribution, and ultimately the N-$C_\alpha$ backbone bond is cleaved, with formation of $c/z$ ions (Figure 1.25).

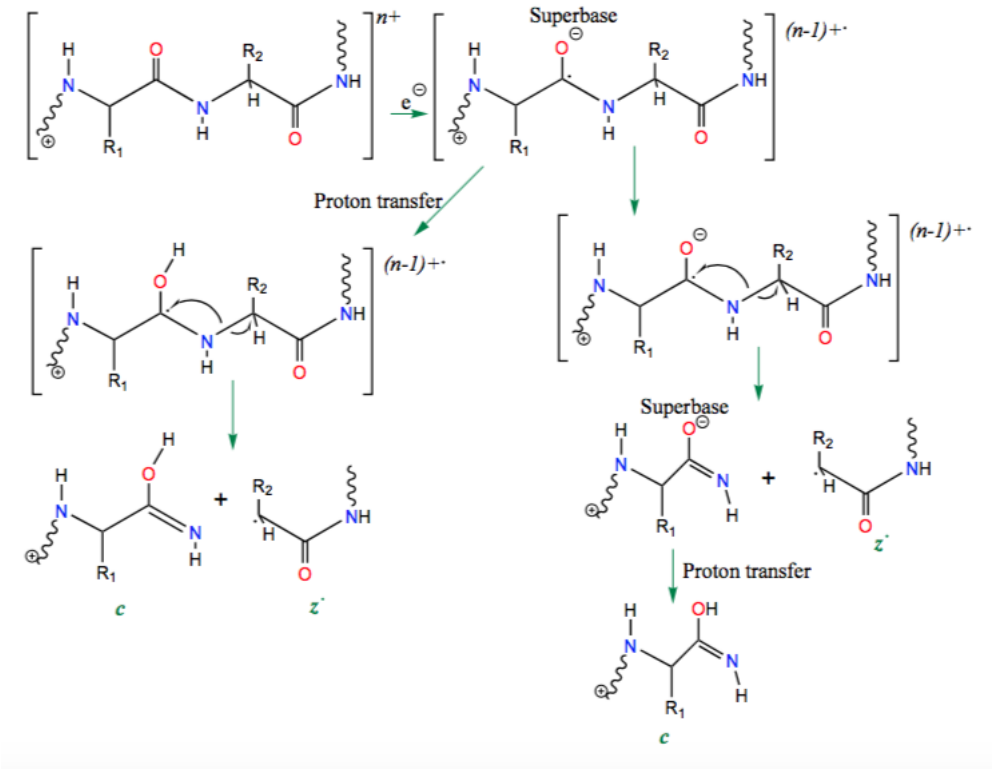



Figure 1.25.: Utah-Washington dissociation mechanism for electron-capture dissociation. It postulates the formation of an aminoketyl radical by capture of an electron into the OCN-amide $\pi^*$orbital, which undergoes proton transfer due to its nature of "super base", forming eventually the charge-reduced species (CRS). The CRS undergoes radical fragmentation with formation of $c/z$ ions. Reproduced with permission from Lopez-Clavijo [106].

A free radical reaction cascade, generated by rearrangement and transfer of the α-carbon radical, can also occur, leading to the formation of ECD/ETD fragment ions with more or less hydrogen atoms than standard $c/z$ ions (*e.g.* $c$-1 ions, or $c^{\bullet}$ ions, and $z$+1 ions, or $z$' ions). The nomenclature for such ions is provided by Zhurov *et al.* [98].

Other electron-based dissociations involve the interaction of singly- or multicharged ions with higher-energy electrons. It is in fact possible to produce an abundant

amount of fragments by irradiating the analyte with electrons at 3-13 eV energy. Because of the higher-energy electrons involved ("hot electrons"), the technique was called hot electron capture dissociation (HECD) [107]. Although different mechanisms and publication-dependent names have been used, electron-based dissociation techniques that use electrons at energy up to 20 eV are enclosed in the name of electron-induced dissociation (EID) [108]. EID is able to produce ECD-like fragments from singly-charged ions.

If high energy electrons are used (>20 eV), the multiply-charged peptides interact with the electrons by becoming radical cations. The radical cations undergo spontaneous fragmentation via side-chain losses, generation of $a/x$ ions, and generation of $c/z$ ions, in a process called electron ionisation dissociation (also abbreviated in EID) [109]. Both the EID techniques mentioned above are characterised by a low fragmentation efficiency of proteins and peptides compared to ECD/ETD, and therefore they are less applied to proteomics studies. Nonetheless, electron-induced dissociation has been successfully used on small singly-protonated molecules [110], polyketydes [111], amino acids [112], small pharmaceuticals [113, 114], and larger molecules [115].

Electron-based dissociation techniques demonstrated several advantages in tandem mass spectrometry for proteomics [97, 98, 100, 101, 102]. The main advantages of low energy ExD's in proteomics include the ability to cleave the backbone bonds of a protein or peptide leaving the more labile post-translational modifications intact, which would be lost with CAD [116]. Furthermore, the more "random" mechanisms involved in ExD make that there are no preferential cleavages for these methods, which are able to cleave deeper in the protein chain than collisional activation techniques [98, 117, 118, 119].

ECD and ETD, whose mechanism involves the formation of a charge-reduced species, are more efficient for highly-charged peptides/proteins (generally 3 or more positive charges). These charge-dependent techniques tend in fact to undergo charge reduction instead of dissociation (electron-transfer no-dissociation, ETnoD) for low charge state ions. ETnoD can also be obtained voluntarily by tuning the electron-transfer dissociation method in order to obtain 1+ peptide ions, which are easier to fragment/assign with non charge-dependent techniques. Especially for doubly-charged ions, it has been observed that some labile interactions are retained with the use of ECD/ETD, decreasing the formation of diagnostic ions. Such ions can be separated by the rest of the amino acid chain by cleaving the intramolecular interaction with supplementary infrared multiphoton dissociation or collisional activations [120, 121, 122, 123, 124, 125]. Such low-energy "heating" techniques are performed in small pulses before or during ECD or ETD, in a procedure called activated-ion ECD/ETD (AI-ECD/ETD). IRMPD-activated ECD/ETD can be specifically called IR-ECD/ETD.

Electron-based techniques targeting negative ions have been developed, such as electron detachment dissociation (EDD) [126, 127, 128, 129], negative-ion electron-capture dissociation (niECD) [130, 131, 132], and negative electron-transfer dissociation (NETD) [133, 134, 135, 136, 137, 138]. All these ExDs target deprotonated species.

### 1.7.4. Photodissociation

Photodissociation is caused by the energization of gas-phase atoms by the absorption of one or more photons [139, 140, 141]. The number of photons needed to achieve photodissociation depends on their wavelength, and it can range from a single high-energy UV-photon to hundreds of IR-photons, which possess less than a tenth of their UV counterparts.

The absorption of multiple low energy IR-photons with subsequent fragmentation is called infrared multiphoton dissociation (IRMPD) [142]. IRMPD is usually achieved by irradiation of the molecules with a $CO_2$ laser, at a wavelength of 10.6 μm. This wavelength is efficiently absorbed by various vibrational modes associated with the C-C, C-N, and P-O bonds of molecules. Because of the low energy of the involved photons (∼0.1 eV) accumulation of multiple photons is necessary to achieve fragmentation, hence the name "multiphoton dissociation" [143]. The low energies involved in IRMPD cause the production of the same kind of ions as CAD [142]. The technique performs optimally with FT-ICR MS instruments, which operate at UHV, thus without the effect of background gases competing with the photodissociation. IRMPD is particularly effective for the analysis of phospopeptides [144, 145, 146].

Ultraviolet photodissociation (UVPD) is diametrically opposite to IRMPD, in a way that it uses a small number of high-energy photons to dissociate molecules. UVPD has been performed with different lasers at different wavelengths, each of them requiring a suitable chromophore for the absorption of UV-photons. Wavelengths of 157 and 193 nm are the most common ways to perform UVPD, achieved respectively with $F_2$ and ArF excimer lasers [147, 148], for photon energies of 7.9 eV/photon and 6.4 eV/photon respectively. Of these two wavelengths, 193 nm has been more prevalent because of the small loss when transmitting through air and the possibility to use common fused silica optics. For UV/vis lasers with photon energies between 3 and 8 eV only one or two photons provide enough energy to excite the molecules at high electronic states. Dissociation might occur directly from the excited states, with formation of a varied array of ions ($a/x$ and $c/z$), or through vibrational redistribution of the internal energy, generating ions more similar to low-energy activation methods (b/$y$ ions). Extensive mechanistic studies performed on UVPD elucidated the photolytic cleavage of the C-$C_\alpha$ bond before the radical elimination to form $a/x$ ions [147, 149, 150, 151, 152, 153].

193 nm UVPD has been successfully coupled with FT-ICR MS, showing excellent performances in top-down proteomics, by generating a wide array of ions that allow correct protein assignments [154, 155].

The main advantages of the photodissociation techniques are the selectivity and tunability based on the used laser, the short activation periods given by the fast laser pulses (in the order of nanoseconds), and, in case of UVPD, high energy deposition with a consequent great array of fragment ions [139, 140, 141].

Figure 1.26 shows the main cleavages caused by the various dissociation techniques. UVPD is able to produce the three of them when UV-photons are generated at a wavelength of 193 nm.

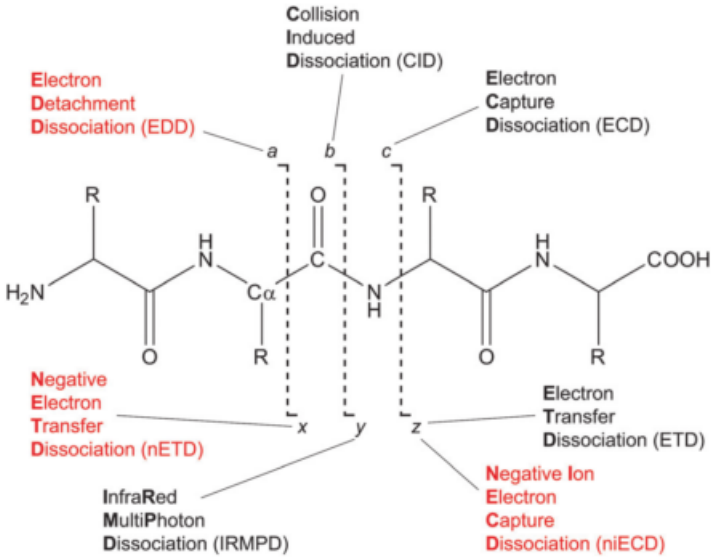



Figure 1.26.: Dissociation of proteins and peptides. ExDs written in black target positive ions, while ExDs written in red target negative ions. Reproduced from Zhurov *et al.* [98].

## 1.8. Separation and identification: hyphenating chromatography with mass spectrometry

Due to its high sensitivity and the ability to deal with complex mixtures, mass spectrometry is the top choice for the analysis of many natural and synthetic compounds. However, when the complexity of the mixture leads to the presence of substances with the same molecular mass, and often the same mass-to-charge ratio, or the presence of several overlapping isotopic distributions which do not allow agile isolation and MS/MS for extended characterisation, a separation method is necessary to obtain simpler mixtures or almost pure samples.

Chromatography is the ultimate separation technique, so that almost every analytical laboratory possesses a chromatographic system. Chromatography is a separation method based on the partition between a mobile and a stationary phase, usually dependent on the polarity of the analyte. The stationary phase is packed into a column, and it is usually made of functionalised silica, while the mobile phase can be an inert gas (as in gas chromatography, GC) or a liquid (for liquid chromatography, LC). The output of chromatography is a chromatograph, a spectrum of the signal corresponding to the different detected analytes as a function of time. For a given method, chromatographic techniques are generally reproducible, allowing a strong correlation between the retention times, the time period that each substance spends eluting in the column, which terminates when the substance is detected, and the properties of the substances, which will come out of the column in a given order, based on their affinities.

Gas chromatography finds its applications in the analysis of complex mixtures of small volatile molecules, from biological samples [156] to petroleum components [157]. It is based to the partition of such molecules between the silica stationary phase and an inert carrier gas, such as nitrogen. In GC, the boiling point of a substance influences its retention time, so that substances that have a lower boiling point, *i.e.* pass faster to the gas phase, will elute first and vice versa. Intuitively, the volatility of a substance is a main factor for its suitability with GC, and proteins are often precluded to this technique.

On the other hand, liquid chromatography bases its separation capabilities on the affinity between a stationary phase of a certain polarity and a solvent or solvent gradient (often water/acetonitrile), which starts from high percentage of aqueous and arrives at high percentage of organic solvent at the end of the chromatographic run. LC has been effective at separating a wide range of compounds, including small molecules [158], polymers [159], and biomolecules [160, 161, 162].

Chromatographic techniques are equipped with detectors able to identify the presence,and sometimes the quantity, previous calibration with internal standards, but not the composition of the analytes. Such detectors include but are not limited to flame ionisation detectors, FID, for GC; or UV/visible detectors for LC. Chromatography can also be hyphenated to mass spectrometry, in a separation/identification technique called GCMS or LCMS.

The hyphenation of liquid chromatography with mass spectrometry, above all with the implementation of ESI and nESI, for nanoscale applications such as nano-liquid chromatography, nLC, and with the excellent performances of FT-ICR MS, brings several advantages to the analysis of proteins and peptides, amongst which the most obvious generation of less complex mass spectra. On the other hand, adding an extra step to the analysis involves technical challenges related to the tuning of the new instrumentation, the necessary solubility of the analytes in the LC gradients,

and longer experimental and analysis times. Nonetheless, LCMS has proved its efficacy in the separation and identification of several mixtures, and it is currently the leading technique for the analysis of complex biological mixtures, including cell lysates, proteins, and peptides [163, 81, 164].

## 1.9. FT-ICR MS in the O'Connor group: the Bruker 12 T solariX FT-ICR mass spectrometer

Research in the physical and analytical chemistry laboratory of Professor Peter B. O'Connor is centered around FT-ICR MS, and especially instrumentation building, instrumentation and method development, and applications of the technique to proteomics and petroleomics. The main research instrument is the 12 T solariX FT-ICR mass spectrometer, showed in Figure 1.27.

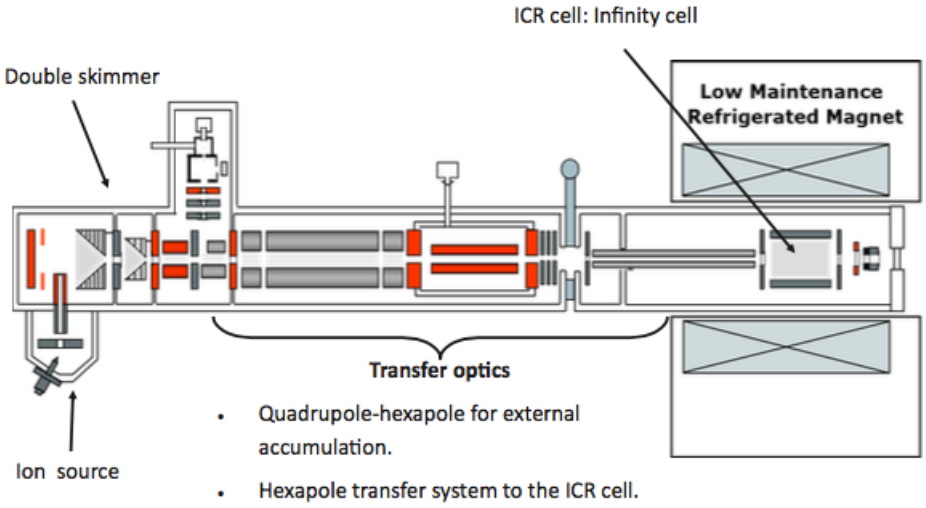



Figure 1.27.: Scheme of the 12 T solariX mass spectrometer. Some of the features are highlighted, such as the orthogonal source; the quadrupole-hexapole system for external accumulation and collision-cell CAD; the Infinity ICR-cell and the superconductive magnet. Adapted with courtesy of Bruker Daltonics, Bremen, Germany.

The solariX FT-ICR mass spectrometer is provided with ESI and MALDI sources, and can be equipped with a variety of ionisation sources, including CI, APPI, APCI, and MAI. It has been further implemented with a custom nESI source, used with the vast majority of proteomics experiments. Furthermore, the instrument is a hybrid Q-FT-ICR MS instrument, allows ion accumulation in the quadrupole section of the instrument, and eventual isolation and fragmentation before transmission to the ICR-cell. The FT-ICR MS/MS features of the Bruker 12 T solariX include nozzle-skimmer in-source dissociation (ISD), CAD, ECD, ETD, EDD, niECD, and EID,

and the group also implemented a 25 W, $CO_2$ laser for IRMPD. The ICR-cell is an Infinity cell.

The instrument is used to challenge the limits of FT-ICR MS itself, both in practice and in theory, in order to achieve the best results possible in terms of data. For the work of research performed for this thesis, the solariX has been used to develop a niche pure mass spectrometry technique called two-dimensional mass spectrometry, and to apply it to TDP and BUP [165], and polymer and glycan analysis [166]. At present, the research performed with the solariX FT-ICR mass spectrometer spaces from characterisation of biomolecules for medical purposes, to protein aging, studying those PTMs that only the high resolving power of FT-ICR MS can detect, such as deamidation [167, 168, 169]; and also cancer research, *e.g.* studying the interactions between anti-cancer drugs and proteins and peptides [170]. The instrument is also used to test novel ionisation sources, and, with the supervision of Doctor Mark P. Barrow, for the analysis of crude oils, fuels, and other petroleum compounds.

A short presentation of the instrument has been given by the author as representative for the whole group for the project *Masses for the masses*, and it is available at http://masses4masses.org.

Figure 1.28 shows the workflow of a typical FT-ICR MS/MS experiment with the Bruker solariX FT-ICR mass spectrometer, and the dissociation techniques most commonly used.

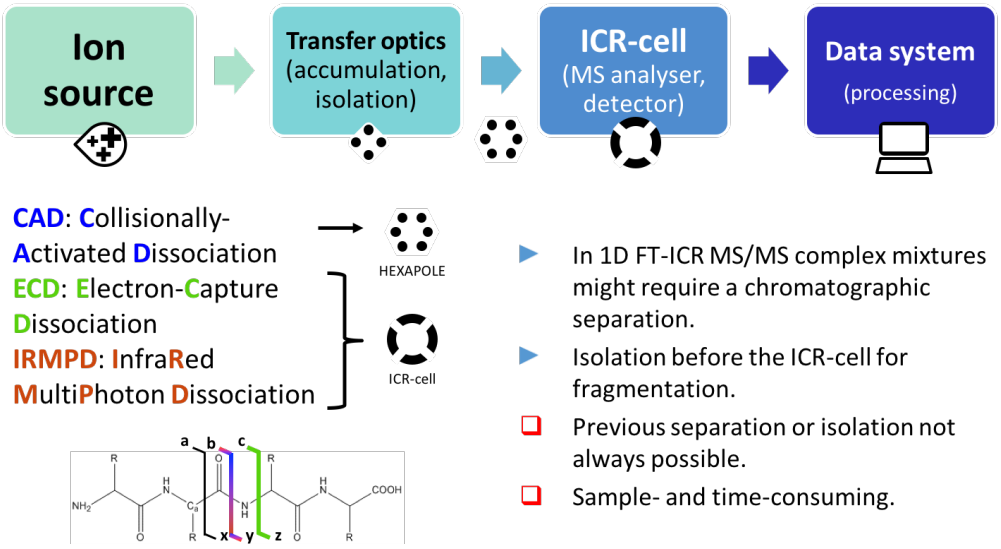



Figure 1.28.: Workflow of a one-dimensional FT-ICR MS/MS experiments with the 12 T Bruker solariX FT-ICR mass spectrometer, with symbols indicating each part of the instrument. CAD is performed by accelerating the analyte ions into a hexapole-based collision cell, while beam-based fragmentations such as ECD and IRMPD are performed inside the ICR-cell. The dissociation techniques are colour-coded with the bond that they break in protein/peptide MS/MS, and the type of ions that they generate (lower left). The bullet points refer to the challenges of the analysis of complex mixtures.

Tandem mass spectrometry is performed by isolating and accumulating an analyte of interest (ionised in the ionisation source) with quadrupolar isolation, and accumulating it in the hexapole collision cell. CAD is performed by accelerating the accumulated ions into the collision cell through the manipulation of the voltages of the extraction lenses situated between a multipole and the other. The generated fragments are subsequently sent to the Infinity cell for mass analysis and detection. ECD and IRMPD are performed through the use of a hollow cathode electrode dispenser and an IR-laser respectively, situated at the end of the instrument, with a direct window to the ICR-cell. For these beam-based techniques, isolation and accumulation of the ion of interest is followed by transfer to the ICR-cell, where the cathode/laser is pulsed, and fragmentation occurs. $MS^n$ can be performed by using in-cell isolation of a primary fragment of interest, and subsequent in-cell fragmentation (with ECD or IRMPD mostly).

The Bruker 12 T solariX FT-ICR mass spectrometer has been used for all the mass spectrometry analysis of this thesis, with the methodology varying for each set of experiments, and described in the respective chapter.

## 1.10. Two-dimensional Fourier transform ion cyclotron resonance mass spectrometry

Hyphenation of mass spectrometry to separation techniques such as chromatography greatly expands the possibilities of the technique towards the analysis of complex mixtures. However, implementation of an additional technique before MS implies an increment in the tuning of experimental settings and the total experimental time. Amongst other factors, the separation timescale in liquid chromatography does not allow for high signal accumulation after extensive fragmentation in the ICR-cell, limiting the final resolution of the resulting FT-ICR MS spectra, and in some cases the visualisation of fragments obtained with low-fragmentation efficiency techniques. A typical example is electron-capture dissociation, whose MS/MS spectra usually need several scans (usually >200) for correct assignments of fragments deriving from entire proteins. Another challenge of complex mixtures relies on the correct isolation of precursor ions before fragmentation in the ICR-cell: isolation of precursors with very similar $m/z$ is in fact very difficult, and even when possible it can cause ion loss, with consequent loss in sensitivity. Furthermore, even when isolating every single precursor of interest in a mixture results possible, it stands to reason that performing such isolations, followed by fragmentation and mass analysis and detection, with respective tunings, would be consistently sample- and time-consuming.

In order to analyse complex mixtures without the necessity to previously separate or isolate precursor ions, a purely mass spectrometric technique was developed, called two-dimensional mass spectrometry, or 2D MS. 2D MS makes use of the principles behind ion motions in FT-ICR MS in order to find a correlation between precursor and fragment ions. Nuclear magnetic resonance (NMR) technology benefitted from a similar implementation when FT-NMR was developed, leading early to two-dimensional (2D) NMR: by influencing the nuclear spins of different atoms with a series of radio-frequency pulses, spectroscopists were able to correlate variables corresponding to those atoms, and to map them in a two-dimensional contour plot. *E.g.* in the nuclear Overhauser effect spectroscopy (NOESY) experiment, the nuclear Overhauser cross relaxation between nuclear spins during the detecting pulse (generally called mixing period) is used to establish the correlations [171]. 2D NMR gives generally more information compared to standard one-dimensional NMR, above all for complex macromolecules.

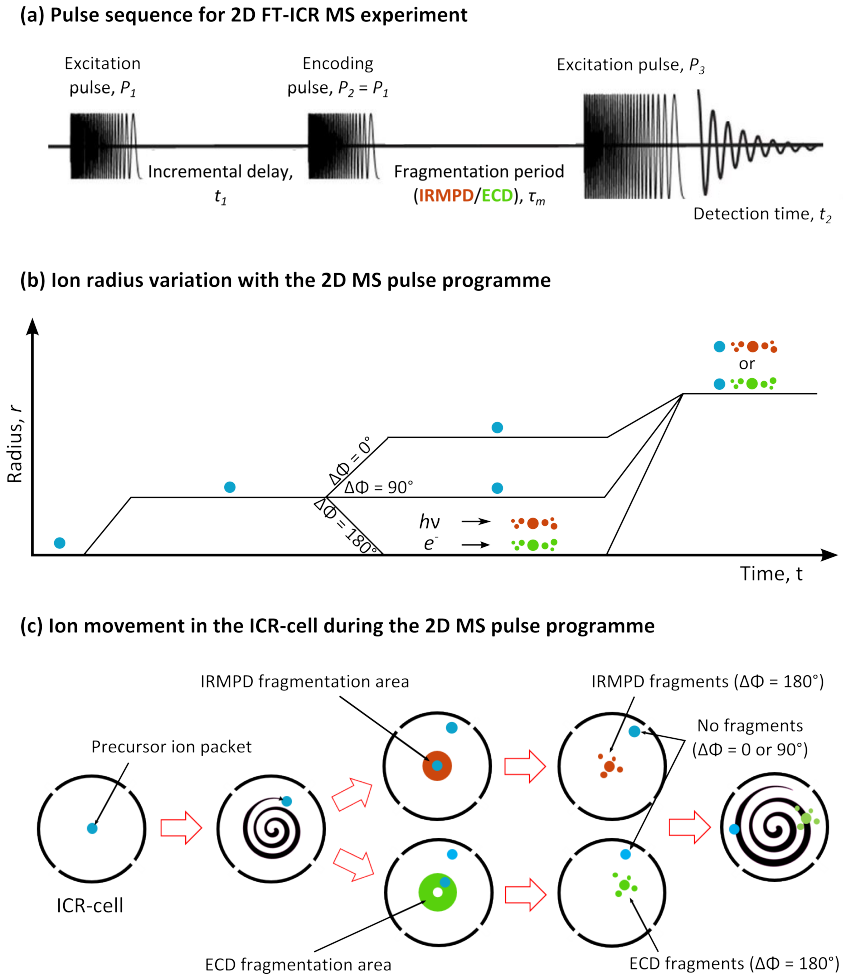



Figure 1.29.: Principles of two-dimensional Fourier transform mass spectrometry. (a) Pulse programme used to obtain the correlation between precursor and fragment ions in 2D FT-ICR MS. (b) Variation of the ion radius int the ICR-cell as a consequence of the different pulses. (c) Movement of the ion packets with the RF-pulses as seen from a transversal section of the ICR-cell. Ions in a cell are constantly orbiting along the magnetic field axis according to their cyclotron motion and are subjected to magnetron motion and oscillations due to the in-cell trapping, but in the present example only the movements driven by the pulses used for 2D MS are represented. Reprinted (adapted) with permission from Floris *et al.* Copyright 2017 American Chemical Society.

Inspired by 2D NMR (and in particular by the NOESY experiment), Pfandler *et al.* designed the first 2D FT-ICR MS experiment [172]. In 2D FT-ICR MS the correlation between precursor and fragment ions is obtained by modulating the ion cyclotron radius using an incremental delay between two identical pulses before fragmentation and detection. The pulse programme used for this experiment is shown in Figure 1.29.a. The ions enter the cell ideally at its centre, and are then excited by an excitation pulse, $P_1$. The excited ion packets rotate inside the cell according to their

cyclotron frequency ($\omega_c$) for an encoding time ($t_1$), to accumulate a phase of $\omega_c t_1$. An encoding pulse ($P_2$), equal to $P_1$, is then applied: this pulse, although identical to the first one, will have a different effect on the ions, now at higher cyclotron radii. As demonstrated by the ion excitation/de-excitation experiment, after a second pulse ions resonantly excited are either excited further or de-excited, depending on their instantaneous phase relative to $P_2$ (Figure 1.29.b). At the end of $P_2$, the ions' cyclotron radii are modulated according to $t_1$ and $\omega_c$. The part of the pulse programme that includes $P_1$, $t_1$ and $P_2$ is called encoding sequence. The encoding sequence is followed by a fragmentation period ($\tau_m$), in which the ions are subjected to a radius-dependent fragmentation, and produce fragment ions whose abundances depend on $t_1$ and the precursor ions' cyclotron frequency ($\omega_c$). Dissociation occurs if after $P_2$ the ions are in the fragmentation area: in the experiment by Pfandler *et al.*, instead of fragmentations, ion-molecule reactions were used to generate $CH_5^+$ by reacting $CH_4^{\bullet+}$ with methane introduced in the ICR-cell. In this case, the product ion is generated with the collisions of the ions excited at higher radii, while the ions de-excited towards the center of the ICR-cell after $P_2$ constitute the remaining unfragmented precursors. As mentioned, in contemporary instruments the introduction of gases in the ICR-cell is avoided, and radius-dependent fragmentations that make use of lasers or electron-dispensers are preferred. In this case, the fragmentation area is at the center of the cell, and the ions that are going to fragment after $P_2$ are the off-resonantly de-excited ions (Figure 1.29.b). After $\tau_m$, $P_3$ excites both precursors and fragment ions, which can then be detected, as per a normal FT-ICR MS experiment, producing a time domain transient ($t_2$). Figure 1.29.c gives an idea of the ion motion in correspondence of the different applied pulses, and of the shape and location of the fragmentation area when using IRMPD or ECD (with a hollow-cathode dispenser). The technique is purely data-independent, because every precursor is fragmented at the same time.

A number $l$ of $t_2$ transients is recorded at incremental values of $t_1$. A two-dimensional (FT/FT) Fourier transform is applied as a function of the mentioned times, along with zero-filling, convolution [173], and apodisation [174]. As a result, a two-dimensional spectrum of the ion cyclotron frequencies is generated, which allows the correlation between the ion cyclotron frequencies of the precursors with those of the fragments. The 2D frequency spectrum then is calibrated and converted into a 2D mass spectrum (Figure 1.30).

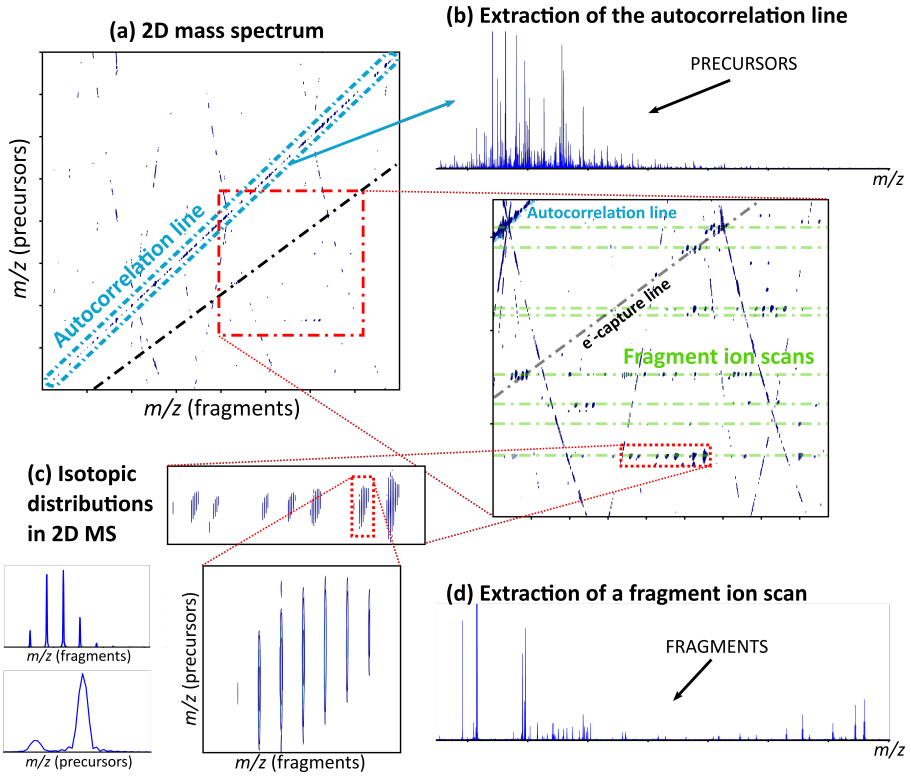



Figure 1.30.: A two-dimensional mass spectrum, output of 2D FT-ICR MS, and various levels of information that can be observed/extracted from it. (a) Two-dimensional mass spectrum seen as a two-dimensional map with the $m/z$ of the precursors in the $y$-axis and the $m/z$ of the fragments in the $x$-axis. The autocorrelation line and an electron-capture line are highlighted. An area is zoomed, to show the horizontal fragment ion scans. (b) Extraction of the autocorrelation line, with generation of a one-dimensional mass spectrum of all the precursors in the mixture. (c) Ion isotopic distributions as seen from the two-dimensional map, and from the vertical and horizontal dimensions. (d) Extraction of a fragment ion scan, showing all the fragments generated by a precursor of interest. See Section 1.10.1 for a detailed explanation on the interpretation of a 2D mass spectrum.

### 1.10.1. Interpretation of a 2D mass spectrum

A 2D mass spectrum is presented as a two-dimensional contour plot reporting on the $y$-axis the mass-to-charge ratio of the precursors, and on the $x$-axis the $m/z$ of the fragment ions (Figure 1.30.a). The $z$ value of such contour plot is represented by the electric signal given by the detection of the ions introduced in the ICR-cell, hence it is related to the intensity of the detected ions. Signals on a 2D mass spectrum ideally appear as circles, which correspond to $m/z$ peaks as seen from a third dimension. In a three-dimensional representation of a 2D mass spectrum, as the one reported in Figure 1.32, the ion intensity could be seen on the $z$-axis, while in a two-dimensional representation it can be deduced by the control lines

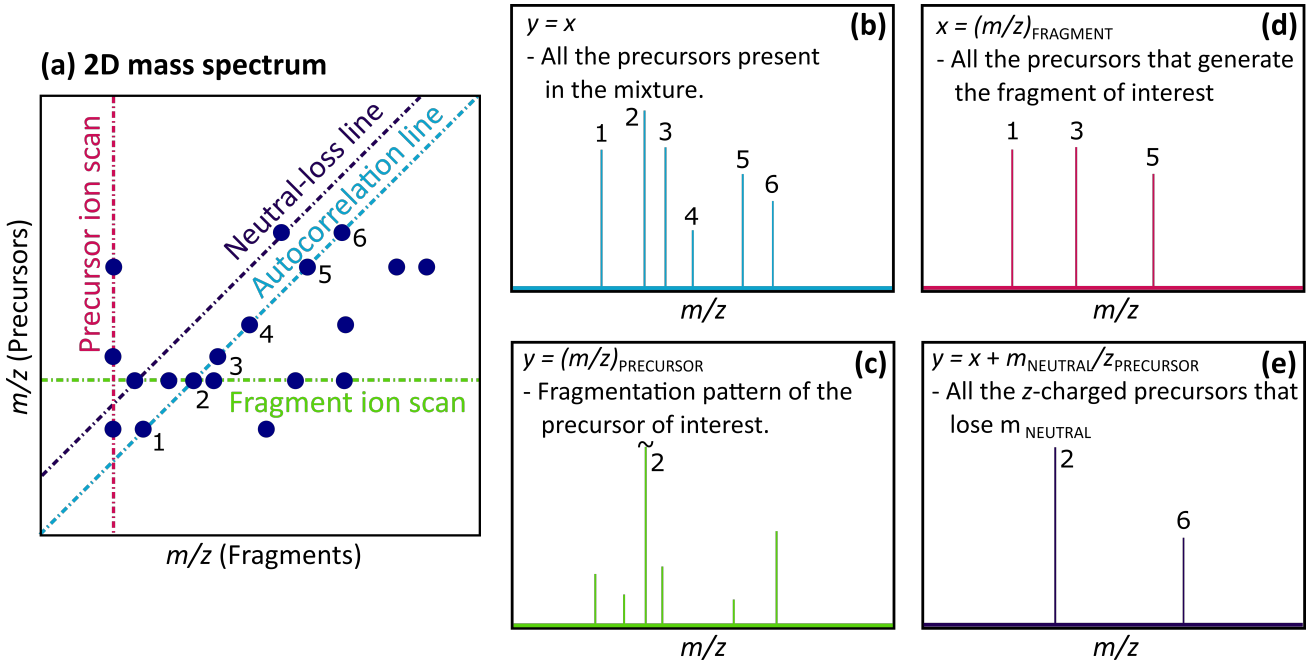



Figure 1.31.: Interpretation of a 2D mass spectrum, with graphical representations of the two-dimensional map (a) and extraction of different lines (b-e). (a) Graphical representation of a 2D mass spectrum, with information about the axis and the main lines of signal used for analysis interpretation. The isotopic distributions for each substance in the spectrum are represented by a single blue dot. (b) Extraction of the autocorrelation line from spectrum a. The generated mass spectrum contains all the precursors present in the mixture, numbered 1 to 6. (c) Extraction of a fragment ion scan. It shows the fragmentation pattern of the precursor 2 in the autocorrelation line. (d) Extraction of a precursor ion scan for a fragment of interest. In this example, the fragment is generated by the precursors 1, 3, and 5. In spectrum a it can be seen that these precursors generate a fragment at the selected $m/z$. (e) Example of neutral-loss line, incident to fragments generated by a specific neutral loss from precursors 2 and 6. It shows the precursors that underwent that neutral loss during the fragmentation perdiod in the ICR-cell. Reprinted with permission from Floris *et al.* Copyright 2017 American Chemical Society.

around the signals corresponding to each peak. As for standard one-dimensional FT-ICR MS, isotopic distributions can be recognised, allowing charge determination, as shown in Figure 1.30.c. The isotopes shown in the figure are elongated in the vertical dimension because of the difference in resolution between the two dimensions. Isotopes are separated in both dimensions, characteristic that allows for additional information during the interpretation of fragmentation patterns (see below).

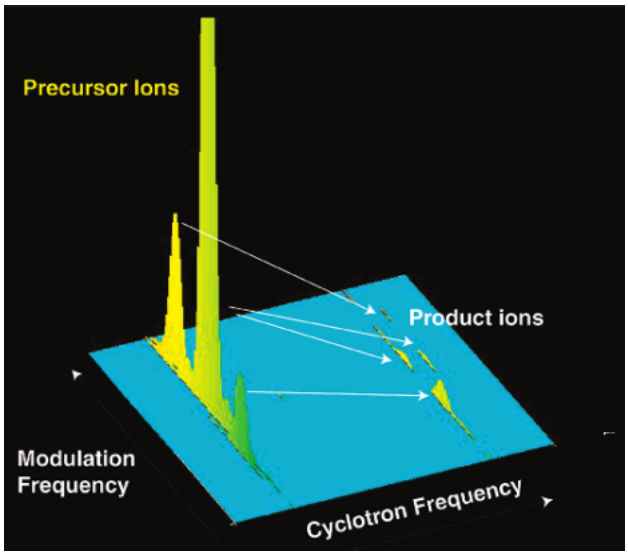



Figure 1.32.: Tridimensional visualization of a 2D mass spectrum, plotted using the cyclotron frequencies of the precursor and fragment ions, as reported by Ross *et al.* [175]. The third dimension is given by the ion intensities.

For data analysis, a 2D mass spectrum can be visualised as a two-dimensional map, and interesting lines can be extracted for a deeper investigation. Extracting a line from the two-dimensional mass spectrum provides a "slice" of the tridimensional structure $(m/z)_{precursors} \times (m/z)_{fragments} \times I$, generating mass spectra more recognisable by the one-dimensional-used scientific community, with intensity $I$ on the $y$-axis, and $m/z$ on the $x$-axis (Figure 1.30.b and d). So far, extracting lines (*i.e.* generating one-dimensional-like mass spectra from the 2D mass spectrum) constituted the only method to efficiently internally calibrate a line on a certain dimension in a 2D mass spectrum, and to provide data usable with modern database search engines.

Interpretation of a 2D mass spectrum relies on the analysis of characteristic lines that can be visualised and extracted from the two-dimensional map. Figure 1.31 reports a graphical example of the main lines of interpretation as seen from the two-dimensional map (1.31.a) and after extraction (1.31.b-e). In the graphical example each isotopic distribution is represented by a single blue dot.

All the peaks along the identity line (*i.e.* the line whose equation is $y = x$) derive from the remaining precursor ion signals in correlation with their own cyclotron

frequency. This "self-correlation" of the precursors inspired the line's name of autocorrelation line. Extracting the autocorrelation line generates a mass spectrum of all the precursors in the mixture (Figure 1.31.b). Horizontal lines in correspondence to the $m/z$ of each precursor $[y = (m/z)_{precursor}]$ will contain signals of the fragments generated by that precursor during the fragmentation period $\tau_m$ in the ICR-cell. Hence, for each precursor, a horizontal ion scan can be extracted. These lines are called fragment ion scans, and generate mass spectra showing the fragmentation patterns of the selected precursor (Figure 1.31.c). Furthermore, for each fragment, it is possible to extract a vertical ion scan $[x = (m/z)_{fragment}]$. Vertical ion scans show signals in correspondence of each precursor correlated to the fragment of interest, reason why they are also called precursor ion scans. Therefore, extracting a precursor ion scan will generate a mass spectrum of all the precursors that generated the selected fragment (Figure 1.31.d). Finally, another kind characteristic line that can be visualised and extracted from the 2D mass spectrum revolves around the loss of neutral molecules or fragments from the precursor ions. In fact, all the signals parallel to the autocorrelation line, with an intercept equal to the mass of a neutral loss divided by the charge of the precursor correspond to the same neutral-loss experienced by the precursors with the selected charge (equation 1.11).

$$y = x + m_{neutral}/z_{precursor} \tag{1.11}$$

These lines of signal are called neutral-loss lines. Extracting a neutral-loss line generates a mass spectrum of all the $z$-charged precursors that underwent a neutral loss of mass $m_{neutral}$ during the fragmentation period in the ICR-cell.

Precursor ion scans and neutral loss lines are prerogative of 2D MS, and the information they provide can be obtained automatically only with this technique.

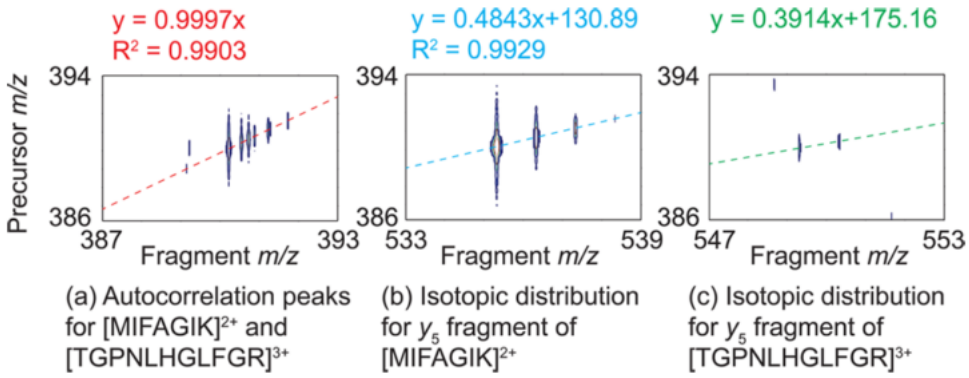



Figure 1.33.: Isotopic distribution of some precursor and fragment ions from an IRMPD 2D mass spectrum of tryptic digested cytochrome c, as seen from the two-dimensional contour plot. Control lines of different colours can be observed for the most intense peaks. The slope of each isotopic distribution is reported. (a) Peaks on on the autocorrelation line corresponding to the peptides $[MIFAGIK]^{2+}$ and $[TGPNLHGLRGR]^{3+}$. The slope of the precursor ions is 1, since they are on the autocorrelation line. (b and c) Isotopic distributions of a $y_5$ fragment respectively from $[MIFAGIK]^{2+}$ and $[TGPNLHGLFGR]^{3+}$. Although the precursors on the autocorrelation line have less than 1Da of difference, the fragments can be correctly assigned thanks to the slope of their isotopic distributions, indicating respectively a 1+ ion generated by a 2+ precursor and a 1+ generated by a 3+ ion. Reprinted (adapted) with permission from van Agthoven *et al.* [176]. Copyright 2015 American Chemical Society.

The nature of a 2D mass spectrum allows to obtain more information from the fragmentation patterns by looking at the isotopic distributions directly from the two dimensional map. *E.g.* a 2+ precursor ion will have an isotopic distribution for which each A+1 isotope would be separated from its neighbour isotopic peaks by 1/2 Daltons per charge. For precursor ions (which are on the autocorrelation line, $y = x$) this is valid in both dimensions (Figure 1.33.a). When the precursor generates fragments, they can be observed on the horizontal line corresponding to the $m/z$ of that precursor [$y = (m/z)_{precursor}$]: each isotope of the precursor will generate fragments, which in the vertical dimension are separated by 1/2 Daltons per charge. However, in the horizontal dimension isotopes of the fragment ions will be separated based on their charge. In case of neutral losses, the charge does not vary, and the slope between the isotopic peaks of the fragment ions is the same as the autocorrelation line, as seen for neutral-loss lines. However, if for example a 1+ fragment is generated by the 2+ precursor ion, on the horizontal dimension its isotopes will be separated by 1 Dalton per charge, while on the vertical dimension the isotopic separations would still be 1/2 $m/z$. The slope between the isotopes in the two-dimensional contour plot is given by the vertical isotopic difference divided by

the horizontal counterpart in $m/z$: $(1/2) : 1 = 1/2$ (Figure1.33.b). For a 1+ fragment generated by a 3+ precursor, the slope would be instead 1/3 (Figure 1.33.c). More generally, for each $z$-charged precursor that generates a $n$-charged fragment by losing a charge $p$ and a mass $m_{neutral}$ during the fragmentation period in the ICR-cell, the fragmentation line (or fragment-loss line) is

$$y = \frac{z-p}{z}x + \frac{m_{neutral}}{z} \tag{1.12}$$

When there is no charge loss, $p = 0$ and the equation is equal to Equation 1.11 for neutral-loss lines.

By observing the slope and the charge of the fragment ion from a 2D mass spectrum it is possible to deduce the charge of the precursor ion. This information allows for more discrimination between precursor ions with similar $m/z$ but different charges, constituting a mean of identification by analysis of the fragmentation patterns even when the vertical resolution would not be enough to properly separate different horizontal ion scans.

The electron-capture line, which can be found in 2D ECD MS spectra, follows a similar equation. The extraction of such line shows all the precursors that "captured" one or more electrons during ECD in the ICR-cell, generating a charge-reduced species. Formally, the equation for the electron-capture line would include the mass of the electron in place of $m_{neutral}$ with an inverted sign (it is an addition and not a mass-loss), but 2D FT-ICR MS spectra acquired so far did not have the necessary resolution to measure such a small mass [177]. Electron-capture lines are approximated to the following equation:

$$y = \frac{z-p}{z}x \tag{1.13}$$

So far electron-capture lines are not of analytical use, but they can be observed to extract information about the involved fragmentation mechanism.

### 1.10.2. Noise in 2D MS

Further lines can be observed in a two-dimensional mass spectrum, constituting noise generated by different sources.

All the signals constituting important information in a 2D mass spectrum are distributed in the straight lines of different directions explained above. However, curved

lines in the two-dimensional contour plot can be also observed. They constitute harmonic lines of the autocorrelation line in the vertical "modulation" dimension, as their modulation is non-sinusoidal. These harmonics are aliased back into the spectrum, and the calibration function causes the lines to curve. Harmonic curves can be reduced with the optimisation of the pulse sequence of 2D MS [178].

Vertical noise "streaks" can be observed in correspondence of the most abundant precursors' $m/z$ (Figure 1.34, top). This source of noise is called "scintillation noise" and originates from the variation in amplitude or frequency of a signal. The name derives from astronomy [179], where the source of noise is given by the atmospheric turbulence that causes the stars to twinkle. It is the equivalent of $t_1$ noise in NMR spectroscopy, caused by various sources of instability [180]. In 2D MS, the most likely cause of scintillation noise is due to the fluctuation in the precursor ion intensities, given by the varying number of ions introduced and transferred to the ICR-cell during the experimental time. Independently from its source, scintillation noise cannot be reduced by accumulating scans as for other sources of experimental noise, because it is proportional to the precursors' signal. Scintillation noise constituted a substantial impairment of the 2D MS technique, due to the difficulty in spectral interpretation in the presence of such noise. Currently, thanks to the use of more and more efficient denoising algorithms, this source of noise no longer represents a significant problem.

The first approach to reduce or eliminate scintillation noise from 2D mass spectra was attempted using an algorithm developed by Cadzow and Wu [181], previously used with success in NMR spectroscopy. The algorithm is based on singular value decomposition (SVD) of the signal into sums to exponentially dumped sinusoids, and makes use of a Toeplitz matrix of the eigenvectors rearranged by decreasing order of their eigenvalues. After an estimation of the noise levels, a number of eigenvectors can be used to reconstruct the entire signal by application to each vertical section of the 2D MS spectrum [182]. The use of algorithms based on SVD becomes less attractive the bigger the data size: for increasing sizes, bigger matrixes are required, with an increase in processing time and required computational capability. To tackle the disadvantages of SVD on 2D mass spectra, whose size is generally $10^3$ times bigger than similar 2D NMR spectra, Delsuc's group developed a denoising algorithm called uncoiling random QR denoising (urQRd) [183]. This algorithm is based on the QR decomposition of a matrix randomly taken from the data set, *i.e.* the decomposition of a matrix in a product of an orthogonal matrix (R) and an upper triangular matrix (Q). In this case, instead of reducing the rank of the matrix containing the SVD components, a randomized low-rank approximation of the matrix that retains more signal than noise (a Hankel matrix) is performed. Curiously, as the developers themselves affirm, they use randomness to reduce noise. The urQRd denoising algorithm has been successfully applied to NMR, FT-ICR MS, and 2D FT-ICR MS data, and it has been demonstrated that it reduces the computational processing time for

denoising of about three orders of magnitude. Furthermore, urQRd allows denoising to be applied to data of virtually infinite size. The maths behind urQRd is described in the research article by Chiron *et al.* [183] and it is not intention of this chapter to further elaborate on its operating principles.

Since its development, the urQRd denoising algorithm has been implemented in the processing program used by the main groups of research on 2D MS. Although the calculations behind this sort of denoising might not be of any concern to MS users, an important parameter has to be taken in consideration when processing two-dimensional spectra, the rank of the Hankel matrix, $K$. The rank $K$ depends on the number of the expected frequencies. However, opposite to what happens with SVD, it is not as restrictive. The smaller the rank, the more noise is removed, but if the randomized matrix is too small it might not be representative of all the signal. Determining $K$ *a priori* is usually quite difficult for systems involving biological/biochemical samples, and particularly for FT-ICR MS/MS or 2D MS, due to the unexpected number of ionised species, hence the unknown number of expected frequencies. Figure 1.34 shows the 2D mass spectrum of tryptic digested calmodulin without (top) and with (bottom) the application of the urQRd denoising algorithm to the data processing. The rank used was $K = 6$.

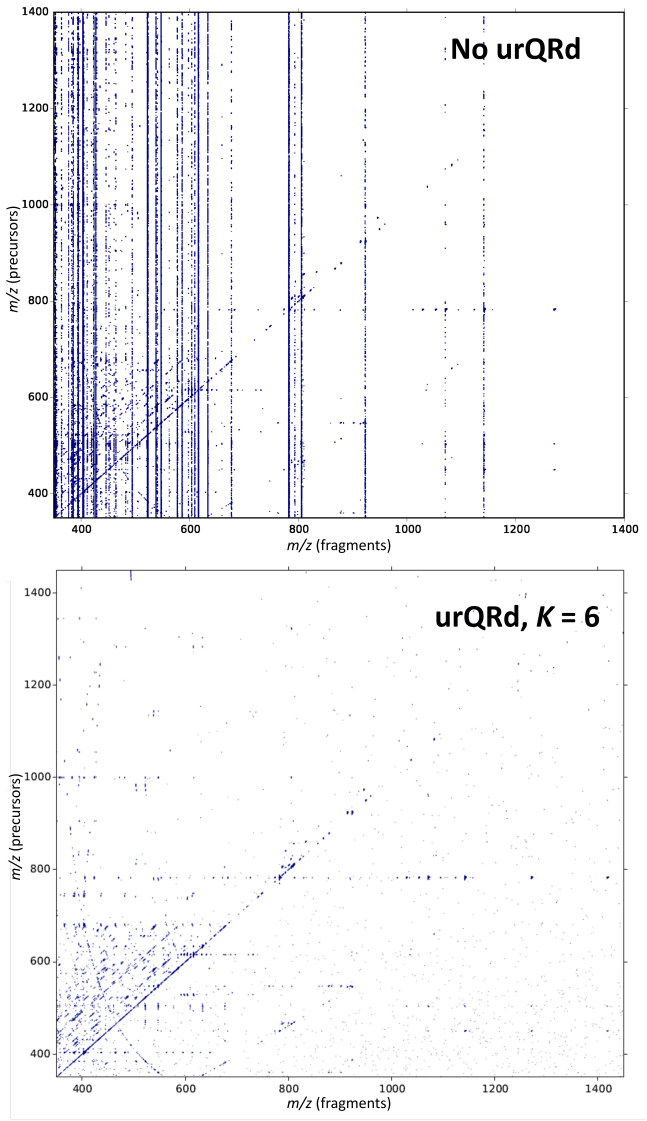



Figure 1.34.: Successful application of the urQRd algorithm to a 2D FT-ICR mass spectrum of tryptic digested calmodulin. In the raw mass spectrum (top) vertical lines of scintillation noise can be observed along the $m/z$ of the most abundant precursors. Scintillation noise is drastically reduced by the application of the urQRd algorithm during processing (bottom). For this spectrum, $K$ was equal to 6.

### 1.10.3. Resolving power and mass accuracy

Resolving power and resolution have been explained for standard 1D FT-ICR MS in Section 1.2.2.

In 2D FT-ICR MS it is possible to define the resolution for both dimensions, vertical and horizontal. Since the $y$-axis shows the $m/z$ of ions in the mixture prior to fragmentation, *i.e.* the precursor ions, the vertical resolution is analogous to the ability to separate two precursor ion peaks. Therefore, since the $x$-axis shows the

$m/z$ of the fragment ions generated by each precursor, the horizontal resolution is related to the separation of narrow fragment ion peaks. Due to the nature of the 2D Fourier transform, the horizontal resolution depends on the number of data points used to acquire $t_2$, while the vertical resolution depends on the number of iterations in $t_1$ (scans). As it can be seen in Figure 1.29.a, $t_2$ is the length of the transient acquired after the fragmentation occurs. Thus, $t_2$ is dictated by the number of data points used for each acquisition and the sampling frequency required, and is easily variable, as for all FTMS acquisitions. $t_1$, as described before, is the delay between the identical initial pulses $P_1$ and $P_2$. The starting value of $t_1 = l_1 \cdot \Delta t_1$ is chosen on the base of the lower $m/z$ of interest, where $\Delta t_1$ is calculated from the sampling rate and the Nyqvist frequency. The number of increments $l_{max}$ is dictated by the vertical resolution needed to effectively separate those precursors during the subsequent 2D-FT-ICR MS experiment.

The inverse proportionality of the resolving power with $m/z$ has been observed for both the vertical and the horizontal dimension [184]. So far, horizontal resolving powers ($RP_h$ or $RP_x$) up to hundreds of thousands ppm have been observed only in top-down proteomics (TDP) studies that did not require a higher vertical resolution to separate the precursor ion peaks, represented by different charge states of the protein under analysis [165]. Vertical resolving powers ($RP_v$ or $RP_y$) are generally much lower, since $l$ is the time-limiting factor of the 2D MS experiment. The difference in resolution between the vertical and the horizontal dimension determines the shape of the signal peaks in the two-dimensional map, justifying the elongation of the circle to a more elliptical shape along the vertical dimension. Since the final size of the datafile is proportional to the number of data points used in the vertical dimension times the number of data points in the horizontal dimension, and since longer transients ($t_2$) and higher numbers of reiterations of the pulse programme ($t_1$) would constitute a longer experiment, the choice of resolution is often led by a compromise (see Section 2.3).

Finally, in 2D MS a total (2D) resolving power is used as an estimation of the entire spectrum, and is calculated as $RP_{2D} = RP_y \times RP_x$.

Mass measurement accuracy (MMA), another point of strength of FT-ICR MS, is a measurement of how the observed masses are close to their theoretical value. It is notoriously defined as $MMA = (m_{obs} - m_{theo})/m_{theo} \cdot 10^6$ in ppm [8], where $m_{obs}$ is the observed mass-to-charge ratio of the peak on the mass spectrum (accurate mass), and $m_{theo}$ its exact mass.

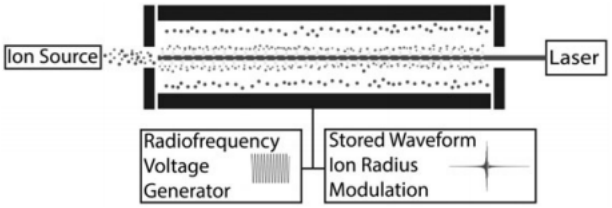
High mass accuracies are obtained with internal calibration of the mass spectra. In 2D MS calibration of the vertical dimension is obtained through the calibration of the autocorrelation line using the theoretical $m/z$ of known precursors. Calibration in the horizontal dimension is performed by internal calibration of a horizontal ion scan (generally with a wide range of fragments spacing through low and high mass-to-

charge ratios), using the theoretical $m/z$ of known fragments. Parameters obtained by internal calibration on a fragment ion scan can be used to calibrate all the fragment ion scans in a 2D mass spectrum, which is the equivalent of an external calibration for the other lines.

### 1.10.4. Alternative 2D MS

The excitation pulses used in the pulse programme for 2D MS are frequency-swept excitation ("chirp") pulses. Another approach to excite ions in the ICR-cell is the stored waveform inverse Fourier transform (SWIFT) [77], which makes use of the inverse FT of the cyclotron frequencies of certain ions to generate a pulse that is able to excite those specific ions. . In 1993, Ross *et al.* used SWIFT to design an alternative pulse sequence for 2D MS, calling the method stored-waveform ion radius modulation (SWIM) [185]. Whereas for the pulse sequence developed by Pfandler *et al.* an incremental delay between two identical chirp pulses is used to achieve modulation of the ion cyclotron radii before fragmentation, SWIM uses one single excitation pulse, after which the cyclotron radius of each ion packet is proportional to the product of the excitation amplitude and its length. The application of $N$ waveforms of known length whose amplitude is varied according to the excitation frequency would modulate the ion cyclotron radii of the precursor ions in the ICR-cell, and therefore the abundance of their fragments after a radius-dependent fragmentation. SWIM 2D MS has been later applied by van der Rest and Marshall for the analysis of oligopeptides [186], and by Ross *et al.* for the analysis of polymeric mixtures and products of combinatorial synthesis used in pharmaceuticals [175]. These studies, respectively in 2001 and 2002, were the first analytical applications of 2D MS on complex mixtures.

2D FT-ICR MS has proven to be a successful analytical technique, but its accessibility is limited by the purchase and maintenance cost of contemporary FT-ICR mass spectrometers. Furthermore, in order to offer the best resolving power amongst the mass spectrometric techniques, the duty cycle of FT-ICR MS/MS is slow, and it is further slowered the more resolution is required. Recently, driven by the curiosity to expand the use of 2D MS on other instruments and by the importance of developing fast data-independent analysis of complex mixtures, van Agthoven and O'Connor studied an *in silico* method to perform 2D MS on a linear ion trap (LIT) [187]. LITs are popular ion manipulation devices [188], to which SWIFT can be reasonably applied. They can generally store more ions than quadrupole ion traps, which is beneficial for 2D MS, since the whole mixture is fragmented/analysed at the same time [189]. van Agthoven and O'Connor simulated a SWIM 2D LIT MS experiment using SIMION, in the configuration shown in Figure 1.35.




**Scheme 1.** Setup for two-dimensional mass spectrometry in a linear ion trap. SWIM pulses are added to the RF drive voltage to preferentially excite some precursors to high radius in the LIT. Precursors remaining at low radius are fragmented by radius-dependent fragmentation methods such as IRMPD, ETD, CID (continuous fragmentation) or UVPD (pulsed fragmentation).

Figure 1.35.: Setup for 2D LIT MS. Reproduced by van Agthoven and O'Connor [187].

The simulation study demonstrated that by modulating the overlap between the precursor ions and the fragmentation area through SWIM, the fragment ion abundances were modulated according to the resonant frequencies of their precursors (or half their encoding frequencies if excitation of the precursor ions leads to ejection). The experiment generated 2D mass spectra similar to the output of 2D FT-ICR MS. In the experiment, the LIT was used as ion manipulation device, and, although it could be used as a mass analyser, the coupling with faster instruments such as time-of-flight (TOF) mass spectrometers or with a drastically higher RP, such as Orbitraps or FT-ICR MS is the most attractive option. In particular, 2D LIT/TOF experiments could require as long as 10 s, much less that the minimum of 20-30 minutes required so far in 2D FT-ICR MS, and compatible with chromatography timescales. The LIT-TOF coupling could lead to online LC-2D MS in which the need of isolation is completely removed.

### 1.10.5. MS$^n$/2D MS

Standard 2D MS gives information equivalent to multiple MS/MS (or MS$^2$) experiments on the analytes of an entire mixture. 2D MS can be implemented to generate information equivalent to MS$^n$ by adding $n$-1 fragmentation steps before the two-dimensional mass spectrometry analysis in the ICR-cell. This procedure is called MS$^n$/2D MS, and it is explained in Figure 1.36 for an extra fragmentation step.

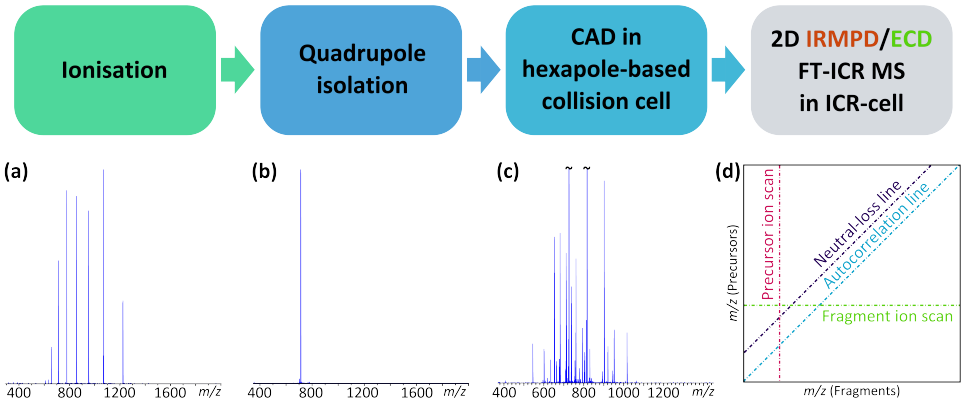



Figure 1.36.: Workflow of MS/2D MS with representative spectra from the different steps of the procedure. Analytes are protonated in the ionisation source, generating a charge-state distribution like the one in insert (**a**). A charge state of interest is then isolated in the quadrupole (**b**) and fragmented with CAD in the hexapole collision cell, generating primary fragment ions shown in (**c**). The generated CAD-fragments are finally sent to the ICR-cell for two-dimensional MS analysis. (**d**) The autocorrelation line (blue) shows the CAD-fragments that entered in the ICR-cell. Horizontal fragment ion spectra (green) show the fragmentation patterns of each precursor ion (further fragments obtained by in-cell fragmentation of CAD-fragments). Vertical precursor ion spectra (magenta) show the precursor ions of each fragment ion. Neutral-loss lines (purple) show the loss of neutrals by precursor ions. Reproduced with permission from Floris *et al.* [190].

The technique shown in Figure 1.36 is an example of MS/2D MS, and it is an experiment that generates information equivalent to MS/MS/MS (or $MS^3$) without any sort of in-cell isolation. The analytes are introduced into the mass spectrometer, and ionised in the ionisation source (1.36.a). A precursor of interest is selected through quadrupolar isolation (1.36.b), accumulated, and fragmented in the hexapole-based collision cell, *e.g.* with CAD (1.36.c). Finally, all the CAD-fragments are sent to the ICR-cell for two-dimensional mass spectrometry with IRMPD or ECD. The output is a two-dimensional mass spectrum containing all the fragmentation patterns of the previously generated CAD-fragments. The technique has been developed by Floris *et al.* to improve the cleavage coverage of TDP 2D MS [190, 191], and made use of 2D MS unique features such as the extraction of precursor ion scans to simultaneously identify specific precursors (*e.g.* all the precursors with the same terminus) or the sequencing of internal ions through the data-independent fragmentation. MS/MS/2D MS (or $MS^2$/2D MS) experiments were also performed on Ubiquitin, using ISD/CAD/ECD (or IRMPD) as fragmentation techniques, with the ISD/CAD/ECD approach obtaining about 87% of the cleavage coverage with a single experiment (see Chapter 4).

MS/2D MS can be used when detailed information of samples of considerable size, such as proteins or protein complexes, is needed.

# 2. 2D FT-ICR MS: two dimensions, infinite possibilities.

The contents of this chapter have been submitted as a book chapter [1].

Principles, background and outputs of 2D FT-ICR MS have been explained in the previous chapter. In this chapter, the acquisition of 2D MS data using contemporary FT-ICR mass spectrometers and processing softwares is explained, and the most recent applications of the technique are detailed. A glossary explaining the terminology in use with 2D MS is also reported.

## 2.1. Data acquisition and processing

Contemporary FT-ICR mass spectrometers allow the users to be in charge of the complete data acquisition, by controlling all the necessary parameters for tandem mass spectrometry analysis, from the ionisation to the detection in the ICR-cell, including accumulation, isolation, and fragmentation, as schematised in Figure 2.1.

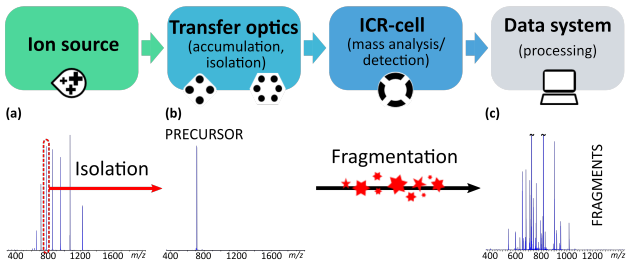



Figure 2.1.: Simplified scheme of a Fourier transform ion cyclotron resonance mass spectrometer and tandem mass spectrometry workflow.

Acquisition of two-dimensional mass spectra is similar in the practice regarding the set of operations that require to be optimised when collecting tandem mass spectrometry data. As for standard FT-ICR MS/MS, the instrument is calibrated through a known, high ionisation-efficiency mixture before data of the real sample are acquired. Ionisation and transfer of the analytes to the ICR-cell are then optimised in order for the user to be able to visualize a full mass spectrum with enough signal-to-noise ratio (S/N or SNR) to discern between real and noise peaks for different substances in the mixture. Other visual parameters such as the peak shape and the oscillation of the peaks around a certain $m/z$ give information about the stability of the sample throughout the process, and are used for a visual confirmation of a correct instrument tuning. A control one-dimensional mass spectrum is usually acquired at this stage, accumulating transients to verify that all the peaks under interest are above the SNR thresholds. If not, the procedure is optimised further, *e.g.* adjusting ion transfer and accumulation. 2D FT-ICR MS analysis requires a stable, continuous source of ions in order to be performed, due to the number $l$ of scans required for the accumulation of a single mass spectrum. For this reason, sources like ESI and nESI, or APPI, able to last for hours depending on the flow rate (and syringe size, for ESI and APPI), have been preferred so far.

When a full mass spectrum is visibly stable on the instrument interface (Figure 2.1.a), the quadrupole-isolation of a suitable peak is optimised, an isolation spectrum is acquired (2.1.b). The isolated substance is then fragmented. The dissociation technique must occur in the ICR-cell and be radius-dependent. Fragmentation is optimised in order to obtain the highest number of fragments, always considering SNR, peak stability, and peak shape. A control tandem mass spectrum is acquired (2.1.c). For

small molecules and peptides, it has been observed that the parameters used for the fragmentation at this stage gave similar results in terms of molecular bonds broken during the 2D MS acquisition, but with differences in fragment abundances [165]. For entire proteins of considerable size, or mixtures so complex that ion isolation is not achievable, the fragmentation is optimised on the entire mixture at once, without isolating. In this case, it is observed how the precursor ion intensity decreases, and the appearance of new peaks in the spectrum, corresponding to fragments of the precursor ions. However, without precursor-ion isolation prior to MS/MS (or the modulation of ion radii of 2D MS) it is not possible to determine which fragment derives from which precursor. Both the mentioned fragmentation optimisations result in one-dimensional mass spectra that will differ from the fragment ion scans extracted from the resulting 2D mass spectrum. This is due to the difference in the number of ions inside the ICR-cell during the "one-dimensional" optimisation and the 2D acquisition, and to the modulation of the precursor ions during the 2D pulse programme, which causes them to oscillate in and out of the fragmentation area. However, the parameters obtained through the 1D-optimisation of the fragmentation efficiency give an idea of what to expect in the 2D mass spectrum, and so far they have been used with excellent results.

Once the parameters for an optimal fragmentation efficiency have been found, the users need to choose the values for the pulses and delays of the 2D MS pulse programme. Figure 1.29.a reports the pulse programme used to achieve correlation between fragment and precursor ions in 2D MS, as explained in Section 1.10.

Studies on the optimisation of the pulse sequence and in particular on the influence of the three excitation pulses on the output of 2D MS, have been performed using IRMPD as the fragmentation technique [178]. Both experimental and simulated data have demonstrated that better two-dimensional spectra are acquired if application of the encoding sequence keeps the fragmenting ions inside the laser radius.

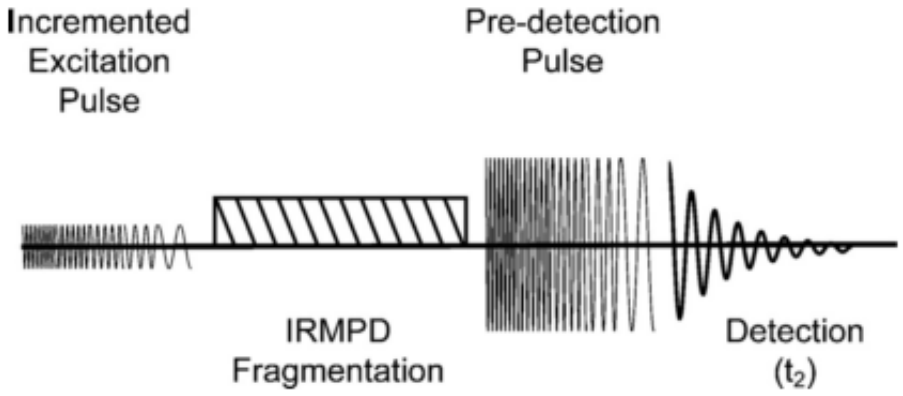



Figure 2.2.: Pulse sequence used to measure the fragmentation efficiency after application of a single excitation pulse, therefore to optimise the 2D MS pulse programme. Reproduced by van Agthoven *et al.* [178].

The effect of a single pulse has been studied by exciting the ions entering the ICR-cell with a chirp of different pulse lengths or amplitudes, and detecting the generated fragments through the usual excitation-detection sequence. Figure 2.2 shows the sequence used for these measurements. The fragmentation efficiency was measured, and the parameters for the best fragmentation efficiency were used to acquire IRMPD 2D mass spectra of the same substances. Experiments confirmed that, by optimising the amplitude and length of the pulses in the encoding sequence, the acquired 2D mass spectra had fewer harmonics and a better SNR. Increasing the observe pulse $P_3$ had different effects on the precursor and fragment ions. The latter see their signal increased, with consequent improvement in SNR, by increasing $P_3$. However, the observe pulse has a different effect on the precursor ions, because their ion intensity in 2D MS depends on the fragmentation in the ICR-cell. The application of an observe pulse which is significantly higher than $P_1$ and $P_2$ overthrows their modulation, inficiating the correlations otherwise allowed in a 2D mass spectrum.

The same sequence used in Figure 2.2 can be used to optimise similar radius-dependent fragmentations for 2D MS.

$t_1$ is back-calculated using the $m/z$ of the lowest precursor of interest, which corresponds to the highest frequency that is going to be observed in the two-dimensional mass spectrum. From the $m/z$ value, knowing the intensity of the magnetic field, it is possible to calculate the digitization rate, and through the Shannon-Nyqvist theorem the sampling rate, which is used to estimate $\Delta t_1$. The lowest precursor of interest can be observed through investigation of the acquired full mass spectrum. The same spectrum gives an idea of the complexity of the mixture, and on the necessary resolving power to be able to discriminate between precursor peaks, hence the number $l$ of reiterations of the pulse programme. $l_{\max}$ corresponds to the number of data points used in the vertical dimension. Two-dimensional bottom-up proteomics

(BUP) studies published so far, involving mixtures of peptides from protein tryptic digestion, used 2048 to 4096 scans. In top-down proteomics (TDP), in which the different charge states of the protein under analysis have usually a significant difference in $m/z$, therefore a much lower resolving power is needed in the vertical dimension, values of $l = 512$ were used.

$P_3$ and $t_2$ are respectively the excitation pulse given to all ions in the ICR-cell in order for them to be detected and the transient length used for the acquisition of the mass spectra, which determines the number of data points used in the horizontal dimension. Both these parameters are chosen during the acquisition of the one-dimensional control mass spectra.

The acquisition of a 2D mass spectrum is then performed, generating a datafile whose size will be proportional to the number of data points used in the vertical dimension times the data points used in the horizontal dimension. As mentioned, these files are generally $10^3$ bigger than 2D-NMR files, and for a long time their data size constituted a limiting factor in the acquisition and processing of two-dimensional mass spectra. *E.g.* the BUP 2D MS spectrum of Figure 1.34 was acquired using 4096 data-points in the vertical dimension and 512k data-points (16-bit) in the horizontal dimension, and the generated datafile before processing was 23.5 GB. The same size was obtained recording the TDP 2D MS spectrum of calmodulin, using 512 data-points in the vertical dimension and 4M in the horizontal dimension. Until the subsequent development of the processing and visualisation programme, SPIKE (see below), and the parallelisation of its code in multiple-core processor units, the acquisition of 2D mass spectra was limited to a size that could be processed and visualised through desktop computers. Therefore, the number of data points in both dimensions were chosen based on the resolving powers involved and the feasibility of the data analysis through office computers. For BUP studies, this affected more the horizontal resolution, reduced from the usual values used routinely in FT-ICR MS/MS (4M data points, 16-bit) to much shorter transients, about 8 to 16 times smaller. A compromise regarding the vertical resolution is also applied when considering the length of the overall experiment: $l$, in fact, takes in consideration the whole duty cycle of a standard FT-ICR MS/MS acquisition. For acquisitions requiring, for example, careful ion transfer or long accumulation times (2-3 s), which have to be added to the pulses, delays, and fragmentation times occurring in the ICR-cell, reiterating the pulse programme 4k to 8k times would result in very long experiments. When other tuning parameters cannot be reduced in order to optimise the experimental time, the number of loops $l$ is chosen based on the minimum necessary resolving power required in the vertical dimension.

The acquired two-dimensional data is processed using SPIKE (Spectrometry Processing Innovative KErnel), a software developed by the University of Strasbourg and CASC4DE (Illkirch-Graffenstaden, France) to adapt the processing of Fourier-

transform based spectroscopies to Big Data [192]. SPIKE is based on a software previously developed for NMR, called NMR processing kernel (NPK) [193]. The program is written in 64-bit Python programming language, and includes features for different FT spectroscopies and urQRd amongst other denoising algorithms. SPIKE is able to interpret the metadata associated to the data-files produced by different manufacturers, and to extract the raw experimental data for its processing. In case of 2D FT-ICR MS, Bruker *Apex* and *Solarix* files can be read. The processing is performed *onfile*, through Hierarchical Data Format, generating HDF5 file format outputs.

The rank $K$ for urQRd, which determines the extent of the denoising of the resulting two-dimensional mass spectrum (see Section 1.10.2), can be selected at this stage. As mentioned, $K$ is difficult to determine *a priori* for biochemical samples in FT-ICR MS, but some suggestions are given based on the experience in the applications of 2D MS to complex mixtures. The work of van Agthoven *et al.* on the comparative BUP study of cytochrome c used ranks of 3 for both the 2D MS spectra acquired. Floris *et al.*, in a BUP mixture of increased complexity used ranks of 6 to 10 [165, 194], and up to 15 when working with polymeric mixtures or with MS/2D MS [166, 190, 191]. $K$ depends on the number of estimated precursor ions, and can be chosen by observing the complexity shown in the one-dimensional control mass spectrum of the entire mixture. If $K$ is too low, the sub-matrix chosen to estimate the total data and exclude noise might not be representative, and signal corresponding to real peaks can be lost. As control for an efficient denoising, the acquisition of a 2D mass spectrum without denoising and one with a temptative value of $K$ can be performed. In order to verify that real signal has not been excluded in the denoising process, all the peaks on the autocorrelation line of the non-denoised spectrum must be on the autocorrelation line of the denoised spectrum. Applying a denoising algorithm to a 2D mass spectrum greatly increases the computational power required and the processing time, although the use of urQRd reduced these delays compared to previous denoising attempts.

The possibility to parallelise the code onto multiple processors greatly reduces the processing time of big data-files, making the option of using processor clusters definitely more attractive than desktop computers. *E.g.*, the processing of the mentioned 23.5 GB, urQRd rank 6, TDP and BUP spectra of calmodulin took more than 12 hours to be processed on a desktop computer with 8 GB of dedicated RAM and 4 core processors, and about 3 hours using 64 processors. The gain in processing speed with parallelisation over multiple processors was reported by Chiron *et al.* [192].

Finally, the latest version of SPIKE, available at https://bitbucket.org/delsuc/spike, was implemented with a novel compression format, which allows fluent visualisation of big processed files on standard office computers.

## 2.2. Data analysis

SPIKE has been equipped with a visualisation feature enabled through the script visu2D.py. A screen of the visualisation program is reported in Figure 2.3.

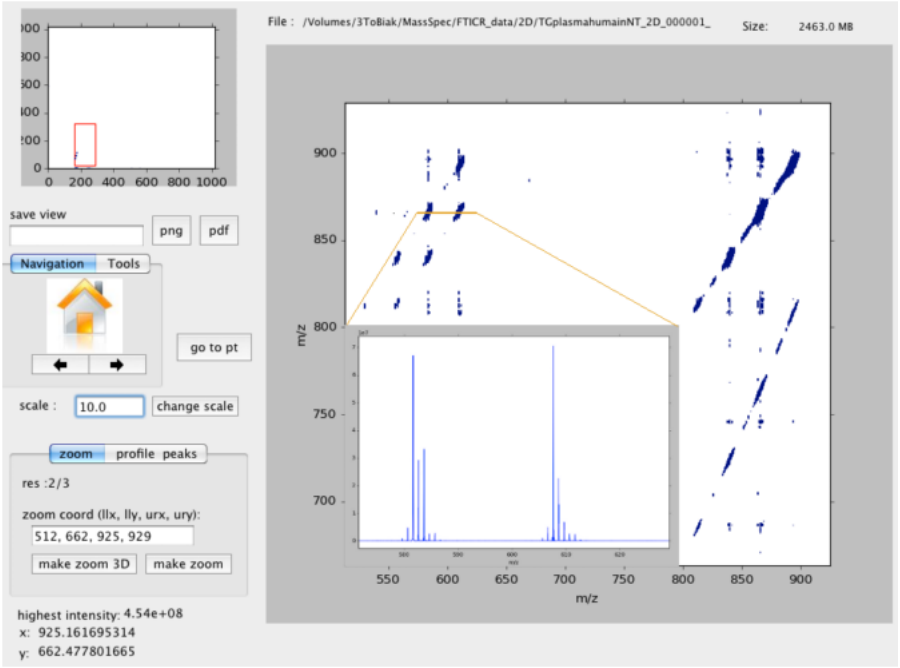



Figure 2.3.: Interface of the visualisation program used in 2D MS, with an extraction of a zoom of the two-dimensional map. Reproduced from Chiron *et al.* [192].

The program creates a custom interface in which the 2D spectrum can be observed by data points or their equivalent in $m/z$. The latter offers the possibility to visualise the two-dimensional contour plot in its common and useful way $[(m/z)_{precursors} \times (m/z)_{fragments}]$. All the visual data analysis is performed using the map in $m/z$ for both the dimensions. The map is visualised at a certain level of scale, representing the depth of the observation of the peaks from "above". The most abundant ions are visualised first, since the intensity of their peaks is higher. The scale can be varied by the user, and increasing the scale increases the number of visualised peaks, up to the noise level. Control lines around the center of the "circles" representing a transversal cut of the peaks from the third dimension are an indication of the peak intensity. The visualisation program has different levels of resolution, and for the visualisation of the complete map the minimum resolution is used. Areas of interest can be explored by zooming on them, by hand, or by using specific $m/z$ coordinates, and the resolution used is higher the deeper is the zoom. The program offers the opportunity to observe the two-dimensional map from the third dimension, including a peak picking program that calculates the centroid of the shown peaks in

both dimensions, or to extract specific lines of signal by inserting their coordinates in $m/z$.

The detailed data analysis of a 2D MS experiment starts from the extraction of the autocorrelation line from the 2D mass spectrum. The spectrum generated by the extraction of the autocorrelation line is a mirror of the mixture before fragmentation, and constitutes information equivalent to the acquisition of a one-dimensional full mass spectrum. The spectrum can be observed through a different window in the visualisation program (like the zoom presented in Figure 2.3), through which some customisation options are available (changing axis scale, colours, etc), as well as the possibility to save the window as PDF, or the extracted data as a spreadsheet with comma separated values (CSV). The autocorrelation line is then calibrated using the theoretical $m/z$ of known precursors, and all the possible peaks are assigned at first based on the their $m/z$. Calibration of the autocorrelation line occurs through extraction of the mass list as a CSV file, conversion into a peak list, and manual calculation using a calibration equation. For each peak on the autocorrelation line, a fragment ion scan can be extracted. From the two dimensional map it is possible to observe the horizontal lines showing the most efficient fragmentation. Usually, a line with many fragments, spread at different mass-to-charge ratios, is chosen and extracted. The generated mass spectrum is calibrated as described for the autocorrelation line (using the theoretical $m/z$ of known fragments), and then all the possible peaks are assigned based on their $m/z$. The parameters used for the successful calibration of this line can be used to calibrate the entire spectrum in the horizontal dimension, and inserted in the code for the mass list conversion, in order to obtain already calibrated mass lists. Fragment ion scans are extracted for all the precursors in the autocorrelation line, and the generated spectra are analysed, and their peaks assigned based on the fragments' $m/z$. Analysis of the fragment ion scans constitutes information equivalent to tandem mass spectrometry targeted at the selected precursor.

By the analysis of the autocorrelation line and the fragment ion scans, the user has already performed a tandem mass spectrometry study on an entire mixture. Particularly complex mixtures may use the extraction of neutral-loss lines or specific fragmentation lines for a successful assignment of the precursors. As explained in Section **??**, neutral-loss lines can be used to assign entire classes of precursors (with the same charge) associated by the same loss. *E.g.* phosphorylated peptides might lose a $PO_2$ group , or N-glycosylated peptides might lose a mannose unit when subjected to IRMPD. Extraction of neutral-loss lines has the potential to identify all those classes of molecules at once, and might be helpful when assignment of precursor ion peaks based on their $m/z$ leads to many different possibilities with similar mass errors. Analysis of neutral-loss lines has been used by Floris *et al.* to successfully assign all the singly-charged linoleated and oleated species at once in a sample of

Polysorbate 80 [166]. Assignment of the linoleated and oleated species allowed the identification of low intensity peaks with an average of 0.04 Da of difference with the different possible assignments.

Vertical ion scans can be extracted for each fragment of interest. Their advantage compared to the neutral-loss lines is that they do not depend on the charge of the precursor ion. Therefore they have the potential to be used for specific precursor ion assignment whenever the precursor generates the fragment of interest. An example of this use is observed in the MS/2D MS work of Floris *et al.*, where by identification of a $b$ ion, successful assignment of many $b$ ions on the autocorrelation line was performed. This feature is particularly important for TDP *de novo* sequencing, where the assignment of a specific fragment ion (in this case an ion retaining the N-terminus of the protein in analysis) could virtually lead to the identification of all the precursor ions retaining the N-terminus in a single scan. Another use for vertical ion scans has been hypothesized by Floris *et al.* [165] in the TDP study of calmodulin, in which the extraction of the precursor ion scan of ion $y_3$ did not show all the charge states of the protein in denaturing conditions found instead in the autocorrelation line. Floris *et al.* hypothesized a charge-dependent fragmentation for which the protein under analysis folded differently based on its charge state, generating different kinds of fragments based on the folding.

Finally, as anticipated in Section 1.10.1, isotopic distribution separation in the two dimensions can be used to obtain information about the charge of their precursors. van Agthoven *et al.* [176] successfully used used this feature to assign fragments of precursors separated by 0.0156 Daltons per charge, based on the slope of the fragments' isotopic distribution, as shown in Figure 1.33.

## 2.3. Applications of 2D MS

After the studies conducted in 2001 and 2002 by van der Rest and Marshall, and by Ross *et al.*, 2D FT-ICR MS was not used anymore until 2010, when it was implemented on a 9.4-T ApexQE FT-ICR mass spectrometer, using a nESI source and IRMPD as fragmentation technique [195]. The analysed samples were small well-known peptides, fragment 1-8 of bradykinin, and substance P. The acquired data were processed with NMR processing kernel (NPK). The obtained mass spectrum, although limited in the vertical resolution by the size of the generated datafile, had an improved total resolution compared to the previous applications of 2D MS due to the use of a fragmentation technique which did not require the injection of gases in the ICR-cell. The experiment revived an interest in the development of 2D FT-ICR MS, which still had to cope with limitations in the processing powers for relatively large datasets, and the impairment given by scintillation noise. Later on, the Cadzow denoising algorithm was developed for use in 2D MS [182], and the technique was

implemented for use with ECD. An ECD 2D mass spectrum of Substance P was successfully acquired, showing features similar to the IRMPD 2D mass spectrum [177]. The technique was also used on a mixture of glycopeptides, sparking the idea to further develop NPK for larger datasets, which at first became NPKV2, and finally SPIKE (described in Section 2.1).

Development of 2D MS for use with different radius-dependent fragmentation techniques, with the use of improved processing and visualisation programs, as well as efficient denoising algorithms to remove the scintillation noise, brought to the generation of high resolution 2D mass spectra, therefore to more analytical applications of the technique.

In 2015, van Agthoven *et al.* used IRMPD 2D MS to differentiate the fragmentation pathways of cholesterol [196]. The experiment was the first example of coupling 2D MS with an APPI source. APPI is a continuous ionisation source for polar and non-polar substances, which generates protonated species and radical cations. Limited in-source fragmentation and oxidation can occur. Cholesterol is a known molecule with a well-described fragmentation, and in this study it was used as a model to demonstrate the features of 2D FT-ICR MS on the analysis of small molecules. Analysis of the APPI-IRMPD 2D mass spectrum of cholesterol is shown in Figure 2.4.

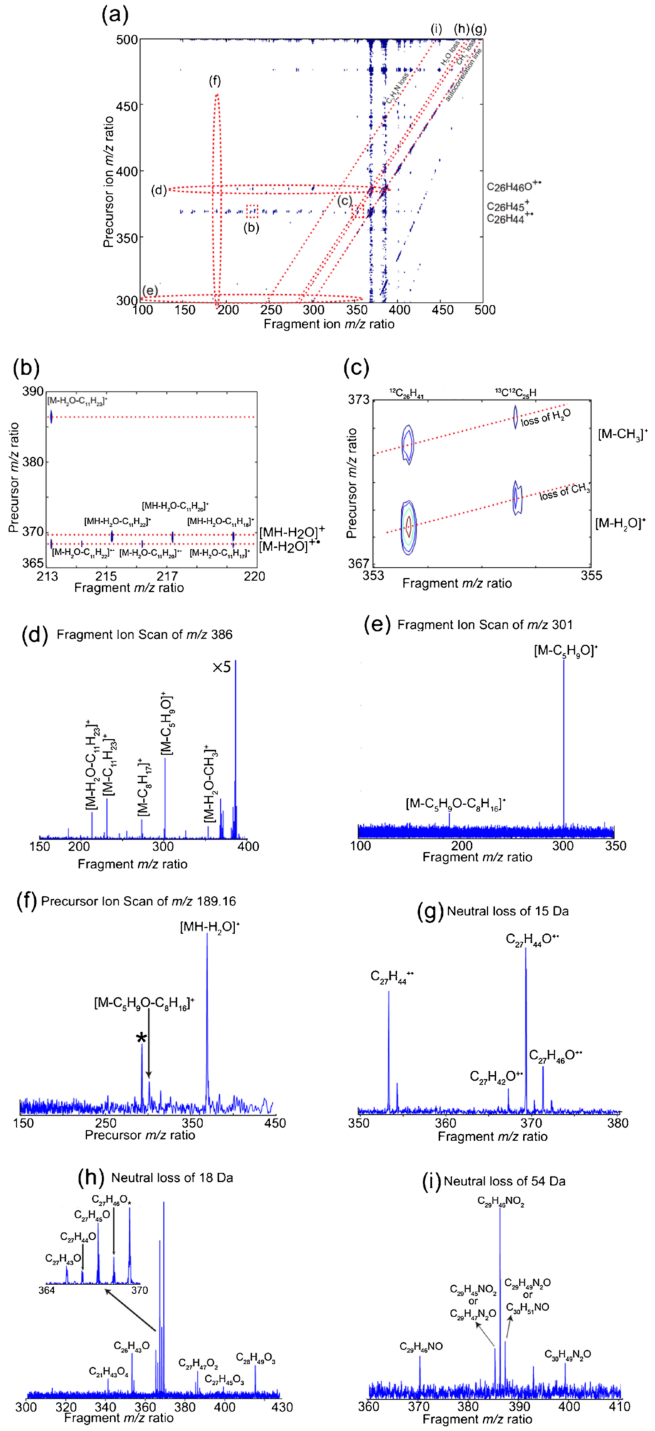



Figure 2.4.: Two-dimensional mass spectrometry analysis of APPI-ionised cholesterol with IRMPD as fragmentation technique. (a) APPI-IRMPD 2D mass spectrum of cholesterol without denoising. Two areas are highlighted in (b) and (c). Two fragment ion scans are highlighted with an ellipse (d and e), as well as a precursor ion scan (f). The autocorrelation line and several neutral loss lines (g, h, i) are indicated by dotted lines. (b, c) Zooms on the 2D mass spectrum, showing area in which respectively fragment ion scans and neutral-loss lines are visible. (d, e) Fragment ion scans of different precursor ions. (g, h, i) Neutral-loss lines showing respectively all the singly-charged precursors that lost $CH_3^{\bullet}$, water, or a nitrogenated group during $\tau_m$. Reproduced from van Agthoven *et al.* [196].

The autocorrelation line was investigated to assign the protonated and radical precursor species, and extraction of fragment ion scans as the examples reported in Figures 2.4.d and 2.4.e allowed for more structural information based on the assigned fragments. The use of 2D MS allowed to separate the fragmentation patterns of very close precursor peaks, such as the ones highlighted at the bottom of the zoom of Figure 2.4.b. The horizontal scans show the fragmentation patterns of the $[M+H-H_2O]^+$ and $[M-H_2O]^{\bullet+}$ species of cholesterol, with a nominal difference of 1 Dalton per charge. Such small mass separation would have been challenging for standard quadrupolar isolation and subsequent fragmentation. The data-independent nature of 2D MS allowed to assign some contaminant peaks generated by reaction of acetonitrile and cholesterol in the APPI source, with production of nitrogenated species. Such species, reported by Marotta *et al.* in previous studies [197, 20], were detected through the extraction of neutral-loss lines of nitrogen-containing species, impossible for cholesterol alone (Fig. 2.4.i). The nitrogenated precursors reported by van Agthoven *et al.* were not particularly abundant, and they would have been missed in a standard MS/MS experiment. However, in 2D MS the fragmentation patterns of all the substances in the mixture are present, allowing the investigation of such substances without further additional experiments. Finally, precursor ion scans were used to obtain more information on the fragmentation mechanisms of some fragments of interest (Fig. 2.4.f).

In the same year, Simon *et al.* started a BUP FT-ICR MS/MS characterisation study on commercially available type I bovine collagen through different fragmentation techniques, CAD, IRMPD, and ECD. Type I bovine collagen is a 406-kDa heterotrimetric triple helix protein [198, 199], whose tryptic digestion leads to hundreds of peptides (681 tryptic peptides *in silico*, assuming no missed cleavages and a variable number of hydroxyprolines) and intuitively to very complicated mass spectra [168]. The full mass spectrum of tryptic digested collagen is reported in Figure 2.5. In such complicated mixture, quadrupolar isolation of precursor ions is extremely challenging without loss of sensitivity.

Although Simon *et al.* reported the successful isolation of some precursor ions, the tandem mass spectra of some species were inficiated by the presence of unassignable peaks, due to co-isolated and fragmented substances. These conditions led to incomplete cleavage coverages of precursor peptides, and unassigned precursors due to the absence of available tandem mass spectrometry. 2D IRMPD MS was used to provide more tandem mass spectrometry information on the entire mixture, by extracting the fragment ion scans of the precursors whose isolation was (a) "contaminated" or (b) not achievable [200]. In the first case, fragment ion scans were used to back up information obtained with 1D-FT-ICR MS/MS, allowing elimination of the contaminant peaks derived from the fragmentation of co-isolated species. For case (b), 2D MS allowed to obtain structural information for the assignment of those

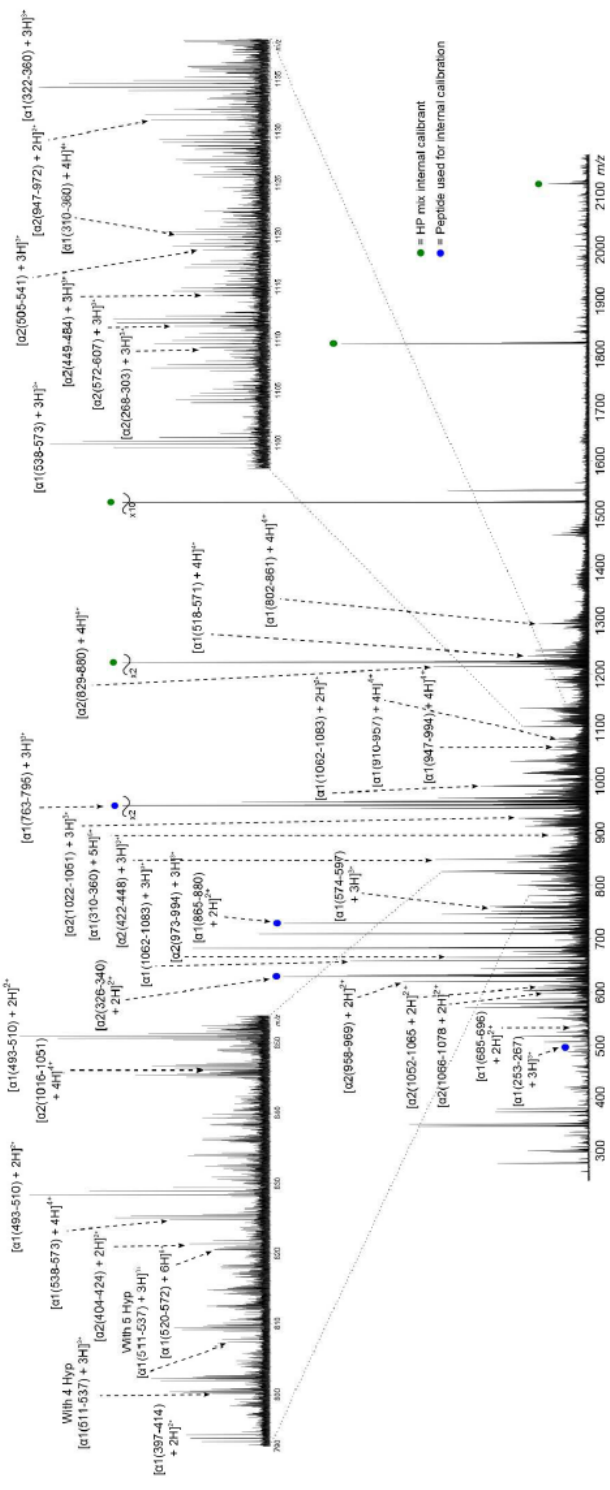



Figure 2.5.: Full mass spectrum of bovine type I tryptic digested collagen, with the labelled assignments. The coloured dots show the peaks used for internal calibration, green for an added high-efficiency polymeric mixture (HP-mix), blue for known peptides. Assignments made on this very dense mass spectrum were possible through the theoretical $m/z$ of the peptide peaks, and with the additional information given by MS/MS and 2D FT-ICR MS. Adapted from Simon *et al.* [200].

peaks, even when only a few cleavages occurred. At that time, 2D FT-ICR MS could not stand alone from 1D MS/MS of complex mixtures, being the technique still in development, but the multidimensional study by Simon, Floris *et al.* represented the first comparative study between MS/MS and 2D MS and a first step towards the analysis of such mixtures.

In 2016, a BUP IRMPD/ECD 2D MS comparative study on cytochrome c demonstrated that 2D MS was able to perform tandem mass spectrometry analysis on complex mixtures, sometimes exceeding the performances of standard MS/MS [176]. Using the unique 2D MS features of separation of masses in two dimensions, van Agthoven *et al.* were able to investigate the fragmentation patterns of precursor ions with a difference of less than 0.02 *m/z*. The cleavage coverage of cytochrome c obtained with two 2D MS experiments of 20 min each was comparable to the one obtained with several IRMPD and ECD MS/MS analyses targeted at the isolated precursors of interests.

The increasing importance of top-down proteomics brought the necessity to develop 2D MS for this mass spectrometry approach to protein analysis. In 2016, Floris *et al.* performed the first two-dimensional TDP study, and compared it to BUP 2D MS of the same protein [165]. The experimental results were also compared against standard TDP and BUP 1D-MS/MS analysis. The work by Floris *et al.* was performed on calmodulin (CaM), a ubiquitous protein able to bind up to four calcium atoms in the human body for signalling purposes [201, 202, 203]. Calmodulin (149 residues, 16.8 kDa) is bigger in size compared to cytochrome c and its structure does not include disulphide bonds. These two characteristics made CaM a suitable model for BUP/TDP 2D MS studies, respectively because it represented a step up in the complexity of the derived tryptic digestion mixture, and because it did not need reduction/alkylation of disulphide bonds, two sample manipulations that often lead to sample losses and further desalting/purification steps before introduction into the mass spectrometer. The study produced the first TDP IRMPD 2D MS of CaM (Figure 2.6.a), showing the fragmentation patterns of all the charge states of the protein at once (2.6.b).

Different charge states subjected to the same fragmentation parameters can produce different fragments. The precursors that generated specific fragments depending on their charge states could be observed through precursor ion scans (Figure 2.6.c). The BUP 2D IRMPD mass spectrum of tryptic digested CaM was also acquired and analysed, and the cleavage coverages obtained with the two-dimensional studies were comparable (23% for TDP and 22% for BUP using IRMPD as fragmentation technique). The study also explored the differences between working with TDP and BUP 2D MS in terms of the vertical and horizontal resolutions required, and the limitations given by the computational capacities in the processing and analysis of mass spectra, due to the generated file sizes of 23.5 GB.

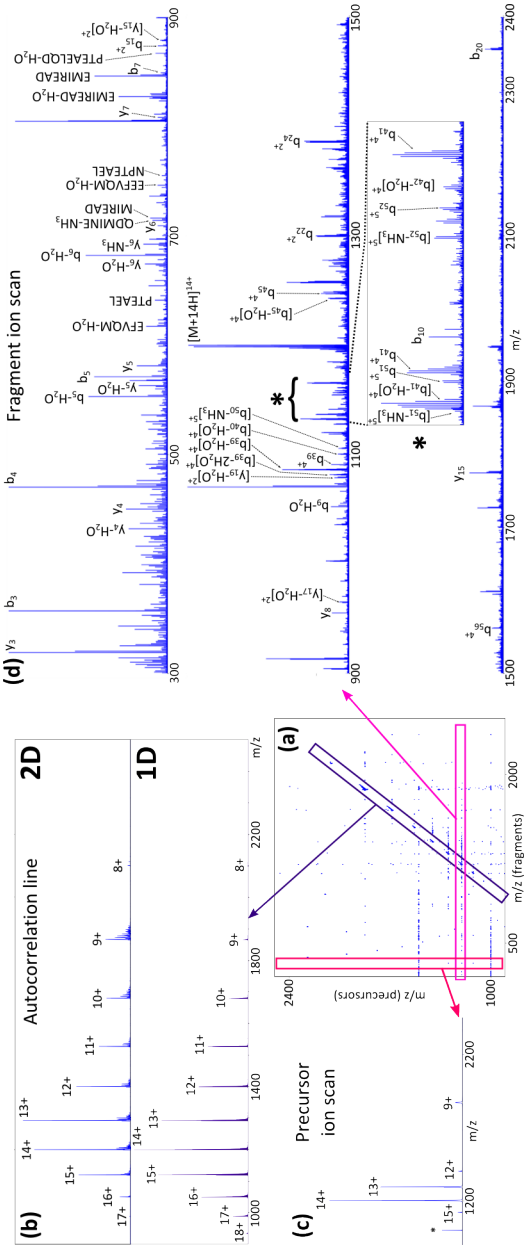



Figure 2.6.: Two-dimensional top-down analysis of calmodulin in denaturing conditions. (a) IRMPD 2D mass spectrum. The autocorrelation line and different horizontal lines of signal can clearly been observed. (b) Extraction of the autocorrelation line (top) and comparison with its homologous full mass spectrum of CaM in denaturing conditions. (c) Precursor ion scan of the ion $y_3$. (d) Fragment ion scan of the 14+ charge state of CaM, showing a complexity of a TDP mass spectrum. Reproduced from Floris *et al.* [165].

The autocorrelation line in Figure 2.6.b (top) is compared to the full 1D mass spectrum of CaM in denaturing conditions (bottom). Such comparisons have been performed for both the studies, comparing the results of 1D MS and MS/MS with 2D MS. Floris *et al.* demonstrated that the cleavage coverages obtained with IRMPD 2D MS were comparable to the 1D IRMPD MS/MS results, with a consistent saving in sample- and time-consumption.

The multidimensional study on calmodulin was later implemented with the use of ECD. By adding a single BUP ECD 2D MS experiment, the cleavage coverage of the protein was increased by more than 20%, confirming the suitability of 2D MS as a fast data-independent analytical technique [194].

Finally, Floris *et al.* applied ECD and IRMPD 2D FT-ICR MS to the analysis of polymeric mixtures, expanding the use of high resolution 2D MS to polymer characterisation [166]. Two-dimensional spectra were acquired for a homopolymer of common use such as poly(methyl methacrylate) (PMMA), and two polymeric mixtures of different complexities, D-α-tocopheryl polyethylene glycol 1000 succinate (TPGS) and Polysorbate 80, commonly used as excipients in pharmaceuticals. The spectra were acquired with IRMPD and ECD as fragmentation techniques, and compared to previous studies on the same substances using MS, or CAD and ECD MS/MS [204, 205, 206]. The multidimensional study demonstrated once again that data obtained with 2D MS with a single experiment were comparable with the results obtained previously with 1D MS/MS studies, establishing the use of 2D MS for polymer analysis. Figure 2.7 reports the two-dimensional analysis of TPGS, and the amount of information extracted from a single IRMPD 2D mass spectrum (2.7.a, b, d-g) in comparison with the some extractions from the ECD 2D spectrum (2.7.c, h-j).

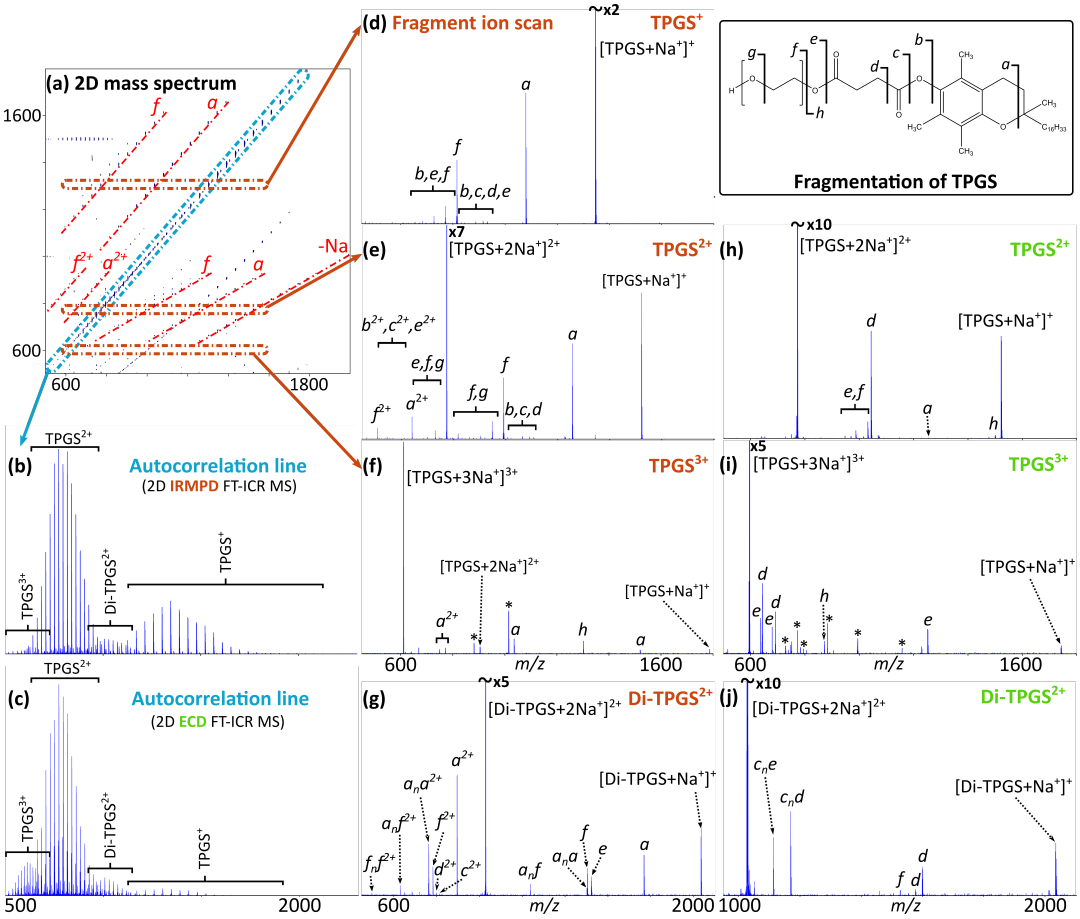



Figure 2.7.: Two-dimensional analysis of TPGS in acetonitrile and $Na_2SO_4$ with IRMPD and ECD as fragmentation techniques, and proposed nomenclature.(a) IRMPD 2D mass spectrum with highlights on the most abundant fragmentations in the mixture. (b) Autocorrelation line extracted from spectrum a. (c) Autocorrelation line extracted from the ECD 2D mass spectrum of TPGS. (d-f) Fragment ion scans extracted from spectrum a, respectively for singly-, doubly- and triply-charged TPGS. (g) Fragment ion scan of Di-TPGS$^{2+}$ extracted from spectrum a. (h, i) Fragment ion scans extracted from the ECD 2D mass spectrum of TPGS, respectively for doubly- and triply-charged TPGS. (J) Fragment ion scan of Di-TPGS$^{2+}$ extracted from the ECD 2D mass spectrum of TPGS. Reproduced from Floris *et al.* [166].

Extraction of neutral-loss lines allowed for the simultaneous detection of the oleated and linoleated species present in the Polysorbate80 mixture, granting the assignment of low-intensity precursors whose possible mass-to-charge ratios differed of 0.04 Da. Polysorbate 80 is a mixture of polymeric synthesis products at very similar $m/z$ [207], nearly impossible to isolate and fragment singularly: the 2D MS analysis of represented the first high resolution tandem mass spectrometry study for the characterisation of such polymeric mixture. Due to the repetitive nature of polymers, preferential fragmentations are repeated along the polymeric distribution of the same

structures, causing strikes of signal clearly visible on the two-dimensional map. These lines (highlighted in red in Figure 2.7.a) are typical of each polymeric mixture, and can be used as a two-dimensional fingerprint for the recognition these samples.

## 2.4. Next steps along the two dimensions

Two-dimensional mass spectrometry underwent several developments in the last decade, to the point of becoming an independent analytical technique with unique features, and constitute a useful platform for complex mixtures in proteomics and polymer analysis. Initial challenges like the scintillation noise, the impossibility to acquire then process and visualise large datasets, and the use and optimization of the technique with different fragmentation methods have now been addressed, bringing to the production of analytically useful two-dimensional mass spectra.

The two-dimensional spectra generated so far, although on top of their resolution for the analytical purposes required, showed resolving powers on the singular dimensions much lower than the RPs achieved routinely in FT-ICR MS. For both the dimensions, this is due to user-imposed limitations on the resolutions in order to save experimental and processing time. On this purpose, Delsuc's and Rolando's groups are working on algorithms that would improve drastically the vertical resolution, based on the non uniform sampling (NUS) on the vertical dimension. Such algorithms have proven to be able to increase the $RP_v$ by several orders of magnitude without prolonging the experiment [208].

The amount of information generated by 2D MS regards the entire mixture under analysis, and at present most of the data analysis is done manually. Tools to automatise some of the data analysis are limited and still in development. In order to tackle the huge amount of MS/MS-equivalent information, it would be useful to automatise the peak-picking process in the two-dimensional map in order to create files readable by the most used database search engines in proteomics. Algorithms for the automatic recognition and assignment of isotopic distributions in the two dimensions, equivalent for example to the SNAP function implemented in Bruker mass spectrometers, would also constitute a matter of taking advantage of all the information available in each single 2D mass spectrum.

Although 2D MS is getting ready to be exported to different mass spectrometry instruments, FT-ICR MS has still improvements to offer. So far, 2D FT-ICR MS has been developed and applied successfully only with the Infinity cell. As mentioned by van Agthoven *et al.* [184], maintaining the ion coherence is very important in 2D FT-ICR MS experiments, due to the FT in the vertical dimension, and the geometry of the cell in this respect can greatly affect the quality of the spectra. Different cells, like the compensated cells on the model of the one developed by Brustkern *et al.*,

or the harmonised cell, and the NADEL cell, mentioned in Section **??**, constitute a promising research ground for further improvements of the technique.

It is hypothesized that coupling of 2D LIT MS with a fast mass analysis technique such as TOF could lead to LC-2D MS experiments in which isolation is eliminated. Such possibility would further open the technique to the analysis of very complex mixtures, eventually enstablishing the technique for routine use in the analytical chemistry laboratory.

## 2.5. Glossary

**Autocorrelation line** Identity line in a 2D mass spectrum ($y = x$), whose extraction generates a mass spectrum of all the precursors in a mixture before their fragmentation. The name derives from the fact that signals along the line are generated by the correlation between the modulation of the precursor intensities with their cyclotron frequencies.

**Electron-capture line** Line that crosses the signals of the charge-reduced species generated by the capture of $p$ electrons during an ECD 2D MS experiment. Its equation is approximated to Equation **??** by ignoring the mass of the electron.

**Encoding pulse** The second pulse in the classical 2D MS pulse sequence, designated $P_2$, identical to $P_1$. After the encoding pulse, the previously resonantly excited ions will be further excited, or de-excited, based on their cyclotron frequency $\omega_c$ and $t_1$.

**Encoding sequence** Part of the 2D MS pulse sequence including $P_1$, $t_1$, and $P_2$. It encodes the cyclotron frequencies of the precursor ions prior to fragmentation.

**Excitation pulse** First pulse of the 2D MS pulse sequence, designated $P_1$. It excites all the ions entering the ICR-cell at higher cyclotron radii.

**Frequency encoding period** Incremental delay between the excitation and the encoding pulse, $t_1 = l_1 \cdot \Delta t_1$, during which the precursor ions orbit according to their cyclotron frequencies. $l$ is the number of increments, and gives the resolution in the vertical dimension. $\Delta t_1$ gives the sampling rate and the Nyqvist frequency. The FT for the vertical "modulation" dimension is calculated along $t_1$.

**Fragmentation period** Fixed period of radius-dependent fragmentation in the ICR-cell, $\tau_m$, between the encoding sequence and the observe pulse.

**Fragment (horizontal) ion scan** Horizontal cross section of the 2D mass spectrum in correspondence of the $m/z$ of a precursor of interest [$y = (m/z)_{precursor}$]. Its extraction generates a mass spectrum of all the fragments generated by that precursor.

**Neutral-loss line** Line parallel to the autocorrelation line, whose extraction generates a mass spectrum of all the precursors of a selected charge that underwent a specific neutral loss (Equation **??**).

**Observe pulse** Last pulse in the 2D MS pulse sequence, designated $P_3$, which excites precursor and fragment ions prior to detection. It is a typical pre-detection chirp pulse applied to bring the ions near to the detection plates for signal acquisition.

**Precursor (vertical) ion scan** Vertical cross section of a 2D mass spectrum, in correspondence of the $m/z$ of a fragment of interest $[x = (\mathrm{m/z})_{fragment}]$ . Extracting a precursor ion scan generates a mass spectrum of all the precursors that generated that fragment.

**Total (2D) resolving power, $RP_{2D}$** Estimate of the resolving power of an entire 2D mass spectrum, obtained by multiplication of the vertical and horizontal resolving powers, $RP_{2D} = RP_y \cdot RP_x$.

**Two-dimensional mass spectrometry, 2D MS** Data-independent tandem mass spectrometry technique that allows direct correlation between precursor and fragment ions, whose output is a mass spectrum in two dimensions. Until 2017 it was a synonym of 2D FT-ICR MS, but recent developments extended the technique to other mass spectrometers.

**Two-dimensional mass spectrum** Two-dimensional contour plot, output of two-dimensional mass spectrometry, reporting on the $y$-axis the $m/z$ of the precursor ions and on the $x$-axis the $m/z$ of the fragment ions. In 2D FT-ICR MS, the vertical and horizontal dimension are the results of a Fourier transform respectively along $t_1$ and $t_2$. In 2D LIT MS the vertical dimension is calculated through a Fourier transform based on the modulation of the precursor ions, but the horizontal dimension depends on the mass analyser used for data acquisition. If an FTMS instrument is used, the obtained 2D mass spectrum still derives from an FT/FT calculation.

**Vertical/horizontal resolving power $RP_{y/x}$** Resolving power along the vertical ($RP_y$ or $RP_v$) or the horizontal dimension ($RP_x$ or $RP_h$).

# 3. 2D FT-ICR MS of Calmodulin: a top-down and bottom-up approach

The content of this chapter has been published in the *Journal of American Society of Mass Spectrometry* in 2016 [165] and in 2017 [194].

Two-dimensional Fourier transform ion cyclotron resonance mass spectrometry (2D FT-ICR MS) allows data independent fragmentation of all ions in a sample and correlation of fragment ions to their precursors through the modulation of precursor ion cyclotron radii prior to fragmentation. Previous results show that implementation of 2D FT-ICR MS with IRMPD and ECD has turned this method into a useful analytical tool. In this work, IRMPD tandem mass spectrometry of Calmodulin (CaM) has been performed both in one-dimensional and two-dimensional FT-ICR MS using a top-down (TDP) and bottom-up approach (BUP). 2D IRMPD FT-ICR MS is used to achieve extensive inter-residue bond cleavage and assignment for CaM, using its unique features for fragment identification in a less time- and sample-consuming experiment than doing the same thing using sequential MS/MS experiments. The study on CaM is then expanded using electron-capture dissociation (ECD) 2D MS as a single complementary bottom-up experiment in order to enhance the cleavage coverage of the protein under analysis. By adding the results of the BUP ECD 2D MS to the IRMPD 2D MS analysis of CaM, the total cleavage coverage increased from ~40% to ~68%.

## 3.1. Introduction

Two-dimensional mass spectrometry is a data-independent tandem mass spectrometry technique that demonstrated its effectiveness in the analysis of small molecules and bottom-up proteomics studies with macromolecules of increasing size [200, 176].In this work, 2D IRMPD FT-ICR MS is developed for use in top-down proteomics, in a comparative study between standard TDP and BUP 1D MS/MS and 2D MS. The method is used for the structural analysis of Calmodulin (CaM) [201, 202], in both TDP and BUP tandem mass spectrometry approaches. Calmodulin is a ubiquitous 16k-Da protein able to bind up to four Calcium(II) atoms in the human body for signalling purposes. CaM has been used as a standard for deamidation [169] and to investigate the behaviour of anti-cancer drugs at the atomic level [209, 170]. Its dimensions and relatively simple structure, given the absence of disulphide bonds, make it the perfect model to perform comparative top-down and bottom-up studies. Finally, the study is implemented with BUP 2D ECD MS of CaM in order to achieve a more extensive inter-residue cleavage coverage of the protein by adding a single 2D MS experiment.

## 3.2. Experimental section

**Materials** Bovine calmodulin (CaM), trypsin (TPCK treated from bovine pancreas), and ammonium bicarbonate ($(NH_4)HCO_3$) were purchased from Sigma Aldrich (Dorset, UK). HPLC grade methanol and formic acid (HAc), were obtained from Fisher Scientific (Loughborough, UK). Water was purified by a Millipore Direct-Q purification system (Merck Millipore, MA, USA).

**Sample preparation** Salt adducts were removed from Calmodulin through 3 kDa molecular weight cut-off centrifugal filters (Amicon Ultra 0.5 mL, Merck Millipore Ltd, Tullagreen, Ireland). CaM was then passed through 0.22 µm centrifugal filters to remove particulates prior to nano-electrospray ionisation. For the top-down approach, Calmodulin (7.5 µM) was dissolved in a 75:25 water/methanol (v/v) solution with 0.3% (v/v) of formic acid. For the Bottom-up approach, Calmodulin (aqueous solution, 40 µM) was digested with trypsin in ammonium bicarbonate 100 mM at an enzyme-to-protein ratio of 1:40. A 3kDa centrifugal filter device was used as a reactor at 37 $^0$C, and the digestion was stopped after 4 hours by isolation from trypsin through centrifugation of the peptides through the molecular weight cut-off filter. The tryptic-digest was diluted to an end concentration of ˜10 µM using a 75:25 water/methanol (v/v) solution and acidified with 0.3% formic acid (v/v).

**Methods** FT-ICR MS was performed on a 12 T Bruker solariX FT-ICR mass spectrometer (Bruker Daltonik GmbH, Bremen, Germany) equipped with a custom nano-electrospray source (nESI). The source used 1.22 mm thin-walled glass capillaries (World Precision Instruments, Hitching, UK) pulled to obtain tips of ~1 µm in diameter with a flaming/brown micropipette puller (Sutter Instrument Co., Novato, California, U.S.A.). The flow rate of the capillaries was ~2 µL/hour in aqueous solution at a capillary voltage of 1000 V. Each experiment used 10-20 µL of sample. Mild in-source dissociation (ISD, 60 V) was used to further remove salt adducts from the nebulised protein and improve the signal of the protonated analytes. IRMPD fragmentation was achieved using a continuous wave, 25 W, $CO_2$ laser (Synrad Inc., Washington, USA) held at 70% of its power output. IR photons were produced at a wavelength of 10.6 µm and pulsed into the ICR cell for 0.3 s prior to detection. ECD was achieved by pulsating electrons from a hollow cathode (1.5 A) for ~0.2 s at a bias of 10 V. The pulse program used for the 2D MS analysis was optimised in order to have the highest fragmentation efficiency according to previous studies [178]. For the TDP 2D FT-ICR MS experiment, 512 scans (increments of $t_1$) of 4M data points were acquired over a mass range of $m/z$ 883.5-3000 $m/z$ on the vertical axis and $m/z$ 147.5-3000 $m/z$ on the horizontal axis, total time of acquisition was ~20 minutes. For the BUP 2D IRMPD FT-ICR MS experiment, 4096 scans of 512k data points were acquired over a mass range of $m/z$ 368.2-3000 $m/z$ on the vertical axis and $m/z$ 147.5-3000 $m/z$ on the horizontal axis; total time of acquisition was ~50 min. For the BUP 2D ECD MS of CaM, 3072 scans of 512 k 16-bit data points were acquired over a mass range of m/z 368.2–3000 $m/z$ on the vertical axis and m/z 147.5–3000 m/z on the horizontal axis; total time of acquisition was ~30 min. 2D MS data were processed with SPIKE [192], using urQRd as a de-noising algorithm [183] ($K = 6$ for TDP 2D MS and $K = 10$ for BUP 2D MS). Data for the 2D MS analysis were processed using a HP desktop computer (Hewlett Packard, Bracknell, UK), via a virtual machine that simulated a Ubuntu 64bit operative system, with 8 GB of dedicated RAM and 4 core processors. One-dimensional IRMPD/FT-ICR MS/MS and ECD/FT-ICR MS/MS spectra of isolated precursor ions were acquired for comparison to the two-dimensional method. All spectra were internally calibrated using known fragment peaks with a quadratic calibration function, and manually interpreted and assigned. For data interpretation of the one-dimensional mass spectra, Data Analysis v4.1 (Bruker Daltonik GmbH, Bremen, Germany) was used. The parameters used for the one-dimensional and two-dimensional mass spectrometric analysis of CaM are provided in the Supporting Information (Table A.8).

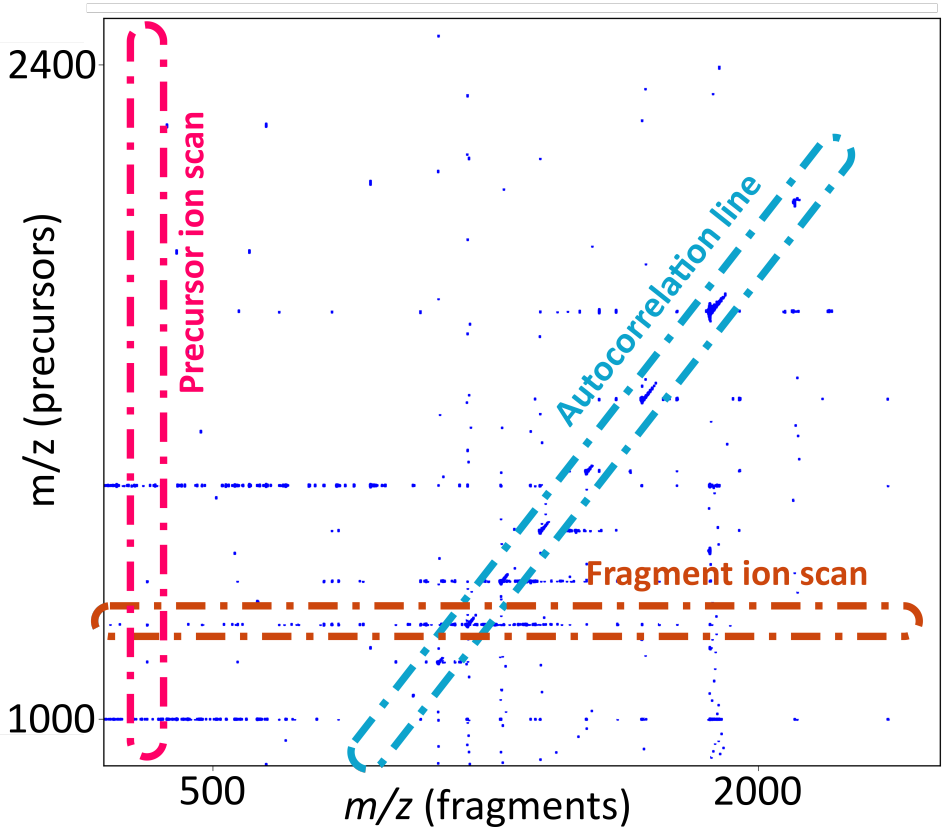



Figure 3.1.: Two-dimensional mass spectrum resulting from the 2D IRMPD MS analysis of CaM in denaturing conditions. The 2D mass spectrum has been acquired with 512 scans of 4M data points over a mass range of $m/z$ 883.5-3000 on the vertical axis and $m/z$ 147.5-3000 on the horizontal axis. Three ion scans, in different dimensions, are highlighted: the autocorrelation line, a precursor ion (vertical) scan, and a fragment ion (horizontal) scan.

## 3.3. 2D IRMPD MS: results

### Top-down Calmodulin

Figure 3.1 shows the TDP 2D IRMPD spectrum of Calmodulin with highlight of the autocorrelation line, a precursor ion (vertical) scan , and a fragment ion (horizontal) scan . The 2D mass spectrum shows a very clear autocorrelation line, representing the correlation of the precursor ions signal with their own cyclotron radius. Figure 3.2 shows the extraction of the autocorrelation line (top) in comparison with a one-dimensional full MS of Calmodulin with the same solution conditions (bottom). CaM shows a wide range of charge states, from 8+ to 17+, as expected under the denaturating conditions used, and that 2D FT-ICR MS is able to detect and fragment them. The signal intensity of the autocorrelation line profile cannot be directly compared with that of the correlating one-dimensional spectrum (Fig. 3.2,

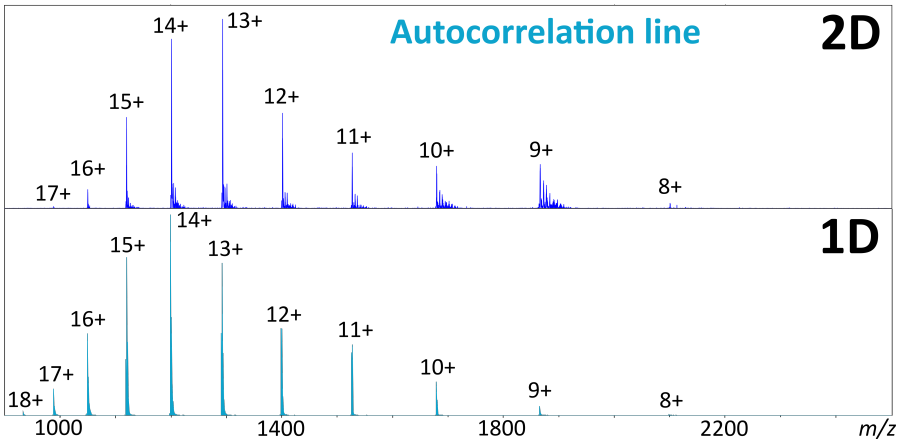



Figure 3.2.: Autocorrelation line extracted from the 2D mass spectrum of Figure 3.1 (top) in comparison with the 1D full mass spectrum of CaM in denaturing conditions (bottom).

bottom, signal averaged for 100 scans) because in 2D FT-ICR MS intensities are not only proportional to the abundance of detected ions; they are also dependent on the change in precursor ion intensity during the fragmentation event. Lower intensity sodium/calcium adducted peaks are also visible in Fig. 3.2, showing an increase in intensity of the sodium ion-adducted peaks in the 2D FT-ICR MS compared to the 1D FT-ICR MS spectrum. As these two spectra were acquired on different days with different samples, it is important not to over-interpret the variation in sodium-adducted peak intensities, but it could indicate there is a greater relative variation (*i.e.* fragmentation) of salt-adducted peaks compared to the purely protonated species. Furthermore, because the 2D FT-ICR MS experiment requires a longer timeframe, it is possible that $Na^+$ leach into solution from the glass capillary, and that would increase the signal of the sodiated species in solution with time. Figure 3.3 shows a vertical precursor ion scan for the $y_3$ ion of CaM, revealing the precursor ions which generate the $y_3$ ion at $m/z$ 319.198. In 2D FT-ICR MS, such a precursor ion scan is simultaneously detected and available for all fragments. Figure 3.4 shows the fragmentation pattern of the 14+ charge state of CaM. It can be seen that the spectrum is very dense, but 2D IRMPD/FT-ICR MS is able to isotopically resolve the fragment ion peaks. The unassigned peaks in spectrum 3.4 may derive from secondary fragments and even "sequence scrambled" ions, as is commonly observed in CAD and/or IRMPD MS/MS [89]. Horizontal profiles have been extracted and analysed for each charge state, giving a cleavage coverage of 23% for the top-down analysis at the specified conditions. Detailed peak assignment tables are included in the Supporting Information (Table A.1). The dynamic range observed for the reported single-scan 2D mass spectrum is ~200. Its value in 1D MS is ~1200, with 200 scans per mass spectrum.

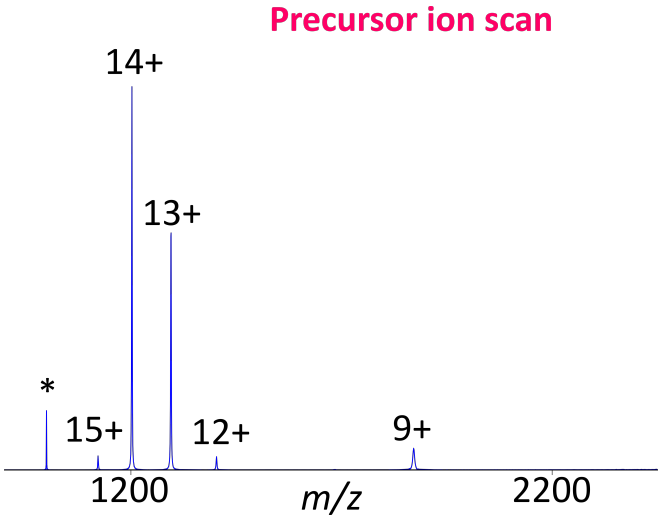



Figure 3.3.: Precursor ion scan of the ion $y_3$. The generated mass spctrum shows all the precurors that generated such ion during the fragmentation period in the ICR-cell.

### Bottom-up Calmodulin

Figure 3.5.a shows the two-dimensional IRMPD-FT-ICR MS spectrum of the tryptic –digest peptide mixture from Calmodulin, with the extraction of the autocorrelation line (3.5.b), a neutral-loss scan (3.5.c), and a fragment ion (horizontal) scan (3.5.d). Figure 3.5.b shows the autocorrelation line profile (top) in comparison with the full one-dimensional (1D) mass spectrum of digested Calmodulin (signal averaged over 100 scans)(bottom). It can be seen from this comparison that 2D FT-ICR MS shows the same fragments present in the digest with few exceptions. A list of the assigned peaks is provided in the Supporting Information for the two-dimensional analysis (Table A.2). Figure 3.5.c shows an example of neutral loss scan, the spectrum shows all the precursors that during fragmentation lose a neutral molecule whose absence would cause a decrease in $m/z$ of 9. According to the neutral loss line equation, $y = x + m_{neutral}/z$, for doubly charged ions, this loss would correspond to loss of a water molecule $[\text{M-H}_2\text{O}]^{2+}$. Figure 3.5.c is therefore showing all the doubly charged ions in the mixture which lose a water molecule during IRMPD fragmentation. A neutral-loss line showing water-loss from doubly charger precursors is chosen as an example the capability of 2D-FT-ICR MS techniques: the same type of profile can be applied to other neutral losses, for example phosphate groups or glycans, respectively from phosphopeptides and glycopeptides, and it constitutes information available readily in 2D FT-ICR MS. Figure 3.5.d shows a fragment ion scan for the ion at $m/z$ 782.380. The horizontal profile (top) is in direct comparison with the corresponding 1D IRMPD FT-ICR MS/MS of the same peptide (bottom). The figures show that results obtained with 2D IRMPD-FT-ICR MS are generally comparable with the one-dimensional approach in terms of fragments shown and sequence coverage, but

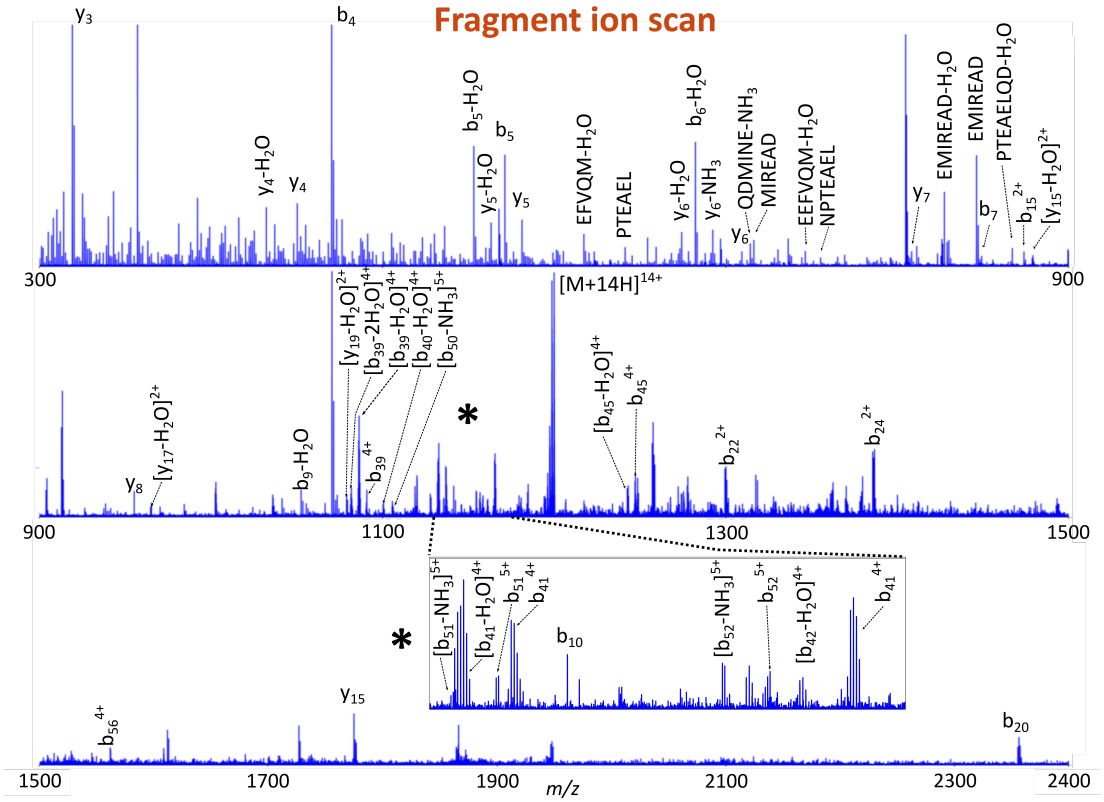



Figure 3.4.: Fragment ion scan showing the fragmentation pattern of the 14+ charge state of CaM. The spectrum is divided in three areas to show all the possible assignments.

do not generate the same signal intensities for the different ions. The differences in ion abundances between the shown fragment ion scan and its homologous MS/MS spectrum derive from the different tuning of the two spectra. Even with the same tuning conditions, although, characteristics such as misalignment of the IR laser, the magnetron expansion of the ion cloud and its own expansion due to modulation in 2D MS make that 2D spectra result different in abundances compared to 1D spectra tuned in a similar way. Furthermore, the parameters for 1D MS are chosen in order to obtain an optimal fragmentation of a single precursor ion at a time, assuming that the ions are constantly inside the laser beam before excitation for detection. In 2D MS, the pulse programme is optimised in a way that the ion packets are modulated in order to be fragmented selectively with the variation of $t_1$, causing their oscillation in and out of the laser beam area, and resulting in different fragmentation pattern intensities. The total sequence coverage of CaM from the detected and fragmented peptide obtained with the 2D IRMPD FT-ICR MS of tryptic-digested CaM was 22% (See Supporting Information tables SA.3 to SA.7). The dynamic range observed for the reported single scan 2D mass spectrum is ~1000. The dynamic range observed in 1D MS for a 200 scans spectrum is ~300.

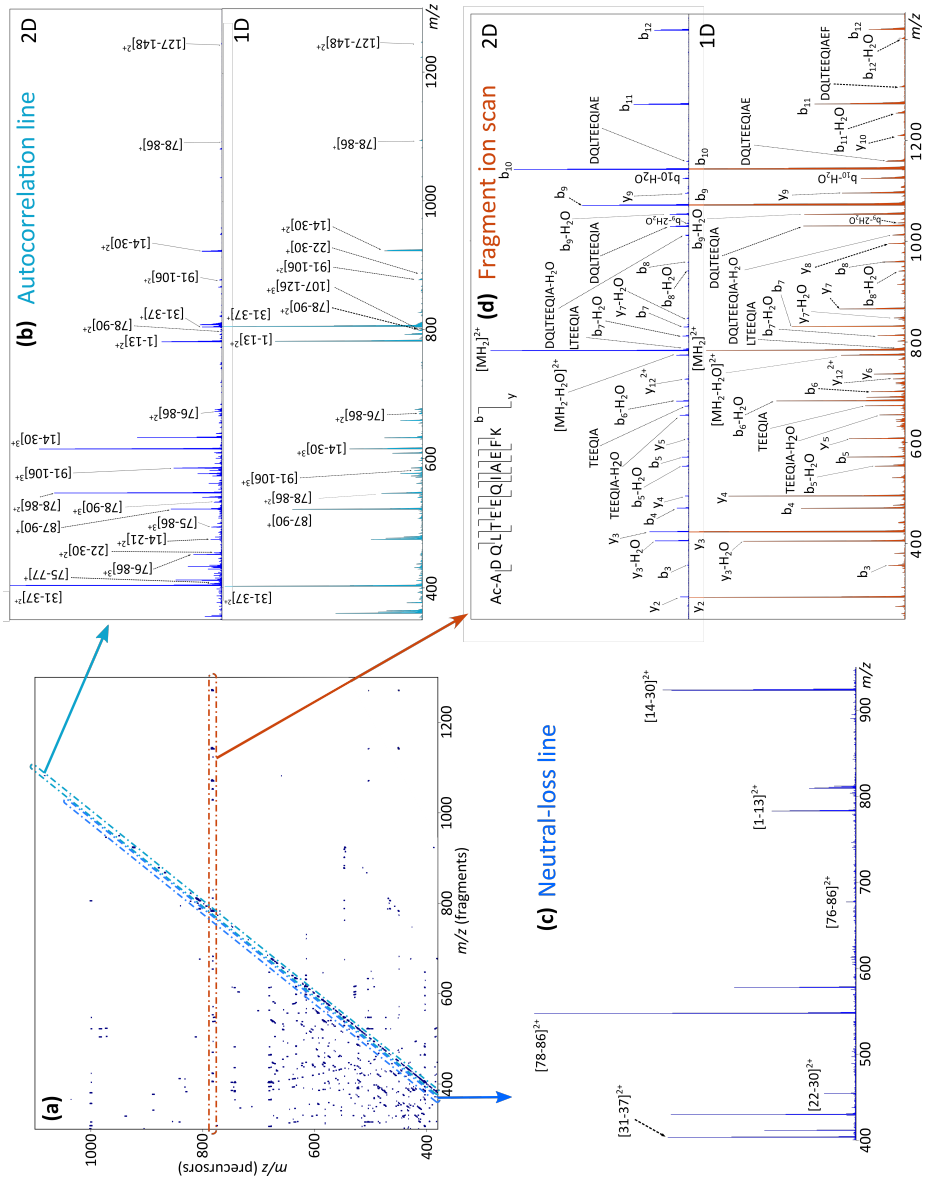



Figure 3.5.: Bottom-up analysis of trypsin-digested CaM. The 2D mass spectrum (**a**) has been acquired with 4096 scans of 512k data points over a mass range of $m/z$ 368.2-3000 on the vertical axis and $m/z$ 147.5-3000 on the horizontal axis. Three ion scans, in different dimensions, are highlighted: **b**, the autocorrelation line, in comparison with the full one dimensional mass spectrum of CaM digest at the same conditions; **c**, a neutral-loss line (information available only in 2D); **d**, a fragment ion (horizontal) scan of the ion $m/z$ 782.380488 in comparison with an IRMPD/FT-ICR MS/MS of the isolated ion with the same $m/z$.

| MS | Cleavage coverage | | |
|---|---|---|---|
| | TDP | BUP | Total |
| 1D MS/MS | 36% | 25% | 52% |
| 2D MS | 23% | 22% | 40% |

Table 3.1.: Cleavage coverages obtained from the 1D IRMPD MS/MS and 2D IRMPD MS analysis of CaM.

## 3.4. 2D IRMPD MS: discussion

2D IRMPD/FT-ICR MS spectra of Calmodulin were acquired using a top-down and a bottom-up approach. The approaches differ greatly with respect to the type and amount of information they provide, and need to be addressed individually when finding the optimal conditions for acquiring 1D and 2D FT-ICR mass spectra.

In every mass spectrometric technique, resolution is influenced amongst other factors by the dataset size during the acquisition. As explained in Chapter 1.10.3, in 2D FT-ICR MS it is possible to define the resolution for both dimensions, vertical and horizontal. The horizontal resolution depends on the number of data points used to acquire $t_2$, while the vertical resolution depends on the number of iterations in $t_1$ (scans). As it can be seen in Figure 1.29, $t_2$ is the length of the transient acquired after the fragmentation occurs. Thus, $t_2$ is dictated by the number of data points used for each acquisition and the sampling frequency required, and is easily variable, as for all FTMS acquisitions. $t_1$, as described before, is the delay between the identical initial pulses $P_1$ and $P_2$. The starting value of $t_1$ is chosen based on the precursor ions of interest, and the number of increments is dictated by the vertical resolution needed to effectively separate those precursors during the subsequent 2D-FT-ICR MS experiment.

The top-down analysis of CaM showed an initial full MS (Figure 3.2, bottom) presenting peaks of different charge states clearly separated in *m/z*, with no overlapping distributions between charge states. Therefore, the vertical resolution needed to be able to separate two narrow peaks in this case would be relatively low, especially compared to the bottom-up procedure (below). For this reason only 512 $t_1$ increments (time steps/data points) were required to ensure the effective separation of precursors. The contrary applies for the horizontal (MS/MS) scans: after IRMPD fragmentation, the density of the spectrum (Figure 3.4) increases greatly compared to the peptide spectra above, requiring a higher horizontal resolution to separate different fragments and provide accurate assignments. In this case, 4M data points were used for the acquired transient, corresponding to a horizontal resolution of 422,473 at 400 *m/z*. Generation of very dense spectra is typical of the top-down approach, which benefits from the resolution and mass accuracy of FT-ICR MS for accurate, resolved assignment. The mass resolving power was measured using the FWHM re-

ported to $m/z$ 400. For the top-down 2D mass spectrum of CaM, the results were: $RP_y$ (vertical) of 182 and an $RP_x$ (horizontal) of 422473.

In the case of the bottom-up procedure, the tryptic-digest solution presents a complex mixture of peptides, whose precursor $m/z$ peaks must be resolved in order to be able to distinguish different precursor peptides and overlapping distributions (Figure 3.5.b). Therefore, a high vertical resolution is required in order to resolve/assign overlapping peaks. In this case, 4096 data points were used for the acquisition in the vertical dimension, 8 times the amount of data points used for the same acquisition in the top-down approach. Since the resolution is proportional to the number of data points used for the acquisition, and that in the vertical dimension this value is determined by the number of $t_1$ time steps, the acquisition of the 2D mass spectrum of digested/complex mixtures generally requires a longer experimental time compared to a top-down experiment due to increased sample complexity. On the other hand, fragmentation of peptides does not generate spectra as dense in peaks as those generated by the fragmentation of entire proteins. As can be seen in Figure 3.5.c, it is possible to resolve the generated peptide fragment peaks with a relatively low horizontal resolution (512k 16-bit datapoints). The mass resolving powers for the bottom-up 2D mass spectrum of CaM were $RP_y = 58766$ and an $RP_x = 1191$ at $m/z$ 400.

Finally it is possible to define a resolution for the two-dimensional map generated in 2D FT-ICR MS, which takes account of the vertical and horizontal dimensions. This composite resolution is the product of the resolution in both dimensions, and is approximately equal to the theoretical peak capacity of the experiment. According to this line of reasoning, the 2D mass resolving powers for the previous spectra are the following: $RP_y \times RP_x = RP_{2D} = 182 \times 422473 \backsimeq 77M$ for the top-down analysis of CaM and $RP_{2D} = 58766 \times 1191 \backsimeq 70M$ for the bottom-up approach, both reported at $m/z$ 400.

The results obtained with 2D FT-ICR MS for both the top-down and the bottom-up analysis of Calmodulin are comparable with standard one-dimensional FT-ICR MS/MS, with a consistent saving in time and sample consumption. For the bottom-up approach, the 2D FT-ICR MS experiment was remarkably faster than a comparative analysis in 1D-MS would be; isolating, fragmenting, and detecting each individual precursor observed in the one-dimensional spectrum of the tryptic digest of CaM could require several hours in comparison. One-dimensional top-down and bottom-up studies on Calmodulin using IRMPD as fragmentation technique were performed for comparison to the 2D IRMPD FT-ICR MS analysis, showing cleavage coverages of the protein up to 36% for the top-down approach, and about 25% for the bottom-up approach, for a total cleavage coverage of 52% (See Supporting Information, Fig. A.1). For the 2D FT-ICR MS experiments the total cleavage coverage obtained with both the approaches reaches ˜40%. The resulting cleavage

coverages for each experiment are reported in Table 3.1. Since precursors have different fragmentation efficiencies, the limitation of using a set fragmentation parameter combination for all charge states is non ideal and could account for the differences between the 2D FT-ICR MS data and individually optimised 1D data. Similar issues arise for LC-MS/MS experiments.

The experiments performed in this work constitute the first example of 2D FT-ICR MS for top-down proteomics, and show that top-down analysis can be performed in 2D MS. In this case, CaM is used as a model for its structure and size, and experiments performed with such a protein constitute the initial step for top-down analysis of complex mixtures. In fact, although the sample used for this analysis was relatively pure, once the procedure is developed it is possible to apply top-down 2D MS on non-pure/complex mixture samples. In such samples, depending on the complexity of the mixture, isolating a protein for standard $MS^n$ analysis might be difficult if not impossible without contamination from ions close in $m/z$. As mentioned, in traditional MS/MS the quality of the spectra depend on the resolution obtained with quadrupole isolation of single precursor ions. Narrowing the isolation window in order to include a single ion (*e.g.* a single charge state of a protein of interest) increases the isolation resolution required, but working with narrow isolation windows in a quadrupole (usually less than 5 Da) inevitably brings substantial signal losses that affect negatively the sensitivity of the final MS/MS spectrum. On the other hand 2D FT-ICR MS provides the fragmentation patterns of all the ions in the mixture without the necessity of quadrupole isolation, preventing the loss of resolution in the precursor dimension even in case of complex mixtures, as it is demonstrated in this thesis for the bottom-up procedure. The resolution in the vertical dimension is higher than resolution typically obtainable with quadruple isolation, and in 2D FT-ICR MS, the vertical resolution is primarily limited by the number of $t_1$ steps (vertical scans) acquired, and by the computational capacity available.

2D FT-ICR MS presents particular computational challenges. 2D IRMPD FT-ICR MS of CaM resulted in files of 23.5 GB for both the approaches, a challenging file size for processing and denoising with standard desktop computers like the one used for this work. 2D NMR, by comparison, usually works with files $10^3$ times smaller [171]. The choice of denoising rank is roughly proportional to the number of precursors in the spectrum, and the lower the rank the more the spectrum is denoised [183]. Denoising a mass spectrum highly increases the data processing time (from about 4 hours to ~12 h for the showed 2D spectra in the described work station system), but it is essential for effective visualisation and interpretation. 2D FT-ICR MS development has been hindered until recently due to limited computational capabilities, and computer technology still represents a significant limitation for the technique, but one that is changing rapidly. The current computational approach to processing this data involves parallelizing the 2DFT, denoising, and visualization code and porting

the data and code up to parallel computing clusters, which will greatly alleviate the current computational bottleneck.

## 3.5. 2D ECD MS: results and discussion

The results of the ECD 2D MS analysis of tryptic-digested calmodulin are shown in Figure 3.6. Figure 3.6.a shows the resulting two-dimensional mass spectrum, with the extraction of the autocorrelation line (3.6.b) and a horizontal fragment ion scan (3.6.c).

The ECD 2D mass spectrum in Figure 3.6.a shows a clear autocorrelation line, and different horizontal lines of signal corresponding to the fragmentation of each precursor. Vertical "noise streaks" are also clearly visible, representing the residual scintillation noise, together with curved lines departing from the bottom left part of the spectrum, constituting harmonic signals. The characteristics of the spectrum are reported in Table 3.2. The vertical resolving power ($RP_y$), the horizontal resolving power ($RP_x$), and the resulting total resolving power ($RP_{2D}$) are reported at *m/z* 400.

| $\mathbf{RP}_y$ | $\mathbf{RP}_x$ | $\mathbf{RP}_{2D}$ | **urQRd rank** | $\mathbf{T}_{acq}$ **[min]** |
|---|---|---|---|---|
| 1k | 60k | 60M | 10 | ~30 |

Table 3.2.: Characteristics of the ECD 2D mass spectrum of tryptic-digested CaM in denaturing conditions.

The autocorrelation line is extracted in Figure 3.6.b. It shows all the species present in the mixture under analysis, derived from the digestion of CaM with trypsin. A total of 19 peaks could be correctly assigned to peptides generated by the tryptic digestion, based on their *m/z* with an average mass accuracy of 0.71±2.86 ppm (mean ± standard deviation, SD). The peaks on the spectrum are present with a range of charges between 2+ and 5+, over a single-scan dynamic range of ~190. The complete peak assignment for the autocorrelation line of Figure 3.6.b is provided in the Supporting Information (Table A.10).

For each precursor, a horizontal ion scan has been extracted, generating a mass spectrum showing the fragmentation pattern of the chosen species. An example is reported in Figure 3.6.c for the ion at *m/z* 782.3805, corresponding to $ADQLTEEQIAEFK^{2+}$, a peptide generated through cleavage by trypsin between the acetylated N-terminal alanine (Ala or A) of the protein and the lysine (Lys or K) in position 13. The same peptide is labelled on the autocorrelation line as $[1\text{-}13]^{2+}$. The fragment ion scan in Figure 3.6.c shows the fragmentation pattern of the peptide, exhibiting a range of *c/z* ions deriving from both the main fragmentation pathways of ECD (generating both $z^{\bullet}$*/z'* and $c^{\bullet}$*/c'* ions), and a minor fragmentation pathway leading to *y* ions. Peaks

corresponding to experimental noise are marked with a * symbol. The spectrum shows cleavages for all the bonds in the peptide. The cleavage coverage is reported on the top right of the figure. The known ECD-fragments for this horizontal mass spectrum have been used to calibrate the entire 2D mass spectrum in the horizontal dimension.

The horizontal scans extracted for each precursor have been analysed, reporting in total 69/148 cleavages for the entire protein, for a resulting cleavage coverage of ~47%. The complete peak assignment for all the fragment ion scans extracted from the ECD 2D mass spectrum of tryptic-digested CaM are provided in the Supporting Information (Tables A.11 to A.17).

As explained above, in 2D MS, analysis of the autocorrelation line constitutes the first means of exploration of the sample, since it constitutes the equivalent in information of a full mass spectrum of the entire mixture. By extracting the autocorrelation line, it is in fact possible to obtain information of what constitutes the mixture before the simultaneous fragmentation of the precursors. In this case, the extraction of the autocorrelation line showed a range of peptides generated by the tryptic digest of calmodulin. In 2D MS, the amplitude of precursor peaks on the autocorrelation line depends both on precursor ion abundance and the decrease of precursor ion abundance caused by fragmentation [178]. In ECD, the efficiency of electron capture increases with the precursor's charge state [210]. As a result, the ion intensities on the autocorrelation line of the 2D ECD mass spectrum are magnified for precursor ions with high charge states compared to a direct infusion mass spectrum. Intensities on the autocorrelation line of the 2D IRMPD mass spectrum of the tryptic digest of calmodulin cannot be correlated to the intensities on the autocorrelation line of the 2D ECD mass spectrum, since the two fragmentation methods have very different mechanisms [176]. The peptides assigned to the autocorrelation line cover almost all of the protein, with a gap between amino acids in positions 31 to 34 (Table A.10 in the Supporting Information).

The initial information gathered with the analysis of the autocorrelation line is improved by analysing the different ion scans (vertical, horizontal, or neutral-loss lines) for each precursor or fragment in the mixture. Although looking at the spectrum of Figure 3.6.a might seem difficult at a first glance, due to the different signals deriving from the scintillation noise and harmonic curved lines, zooming on the two-dimensional contour plot shows clear horizontal lines of signal corresponding to the fragmentation patterns of the precursors. An example is reported in Figure 3.7, showing an expansion of the ECD 2D mass spectrum. Isotopic distributions of different ECD-fragments with different charges can be observed, deriving from a 2+, a 3+ and a 4+ precursor. Such isotopic patterns result more elongated in the vertical dimension because of the difference in resolution between the two dimensions. In this work, 3072 scans were enough to obtain a resolution for which the precursor

ion peaks could be separated, while the number of 16-bit data points used for the acquisition in the horizontal dimension was 512k. 512k data points is generally much lower than the number of data points used routinely in FT-ICR MS in order to obtain the best resolving powers in the mass spectrometry field, but it is still ~170 times higher than the number of data points used in the vertical dimension. 2D MS experiments generate files that contain all the MS/MS information in a single spectrum; files generally $10^3$ times bigger than the average 2D NMR spectrum, to make a comparison with a similar technique. In this work, the horizontal resolution is kept "low" (although giving a horizontal resolving power $<10^4$) in order to generate a file which is easily manageable by a desktop computer. Furthermore, the vertical resolution is kept to a minimum in order to speed up the overall experiment, obtaining the necessary resolving power to separate all the precursors in the vertical dimension (shown in the autocorrelation line) with an experimental time of about 30 min.

The line labelled as $[1\text{-}13]^{2+}$ in Figure 3.7 is a section of the fragment ion scan of Figure 3.6.c as seen from the two-dimensional mass spectrum. Analysis of the horizontal ion scans of the fragmenting analytes improved the exploration of the coverage obtained with IRMPD 2D MS studies on calmodulin. The use of different fragmentation techniques is known to bring complementary information to the structural analysis of proteins, above all when the techniques generate different kinds of fragment ions, such as IRMPD and ECD. The 2D ECD mass spectrum and the 2D IRMPD mass spectrum of the tryptic digest of calmodulin yielded good cleavage coverage for different tryptic peptides, thus resulting in a high cleavage coverage. As a result, combining 2D ECD and 2D IRMPD analysis can be a powerful tool for bottom-up proteomics and can be used to good advantage in order to locate and identify post-translational modifications or point mutations. In fact, although similarities can be observed in the cleaved bonds for the two techniques, the 2D MS experiment performed with ECD enhances the cleavage coverage of the whole protein by fragmenting extensively ions that showed only multiple water losses with IRMPD 2D MS. However, as for the IRMPD study, not all the precursors showed significant fragmentation, because using a set fragmentation parameter combination is not ideal for all charge states.

Finally, Figure 3.7 (bottom) shows the cleavage coverage map of CaM obtained with IRMPD (orange) and ECD (green) 2D MS. By summing the information obtained with BUP 2D ECD MS and the bottom-up and top-down 2D MS analysis of calmodulin, the cleavage coverage of the protein has risen to ~68%.

## 3.6. Conclusions for Chapter 3

2D IRMPD/FT-ICR MS has been used for the top-down and bottom-up analysis of Calmodulin, and fragmentation patterns of CaM have been obtained with both approaches. The results obtained are comparable with one-dimensional FT-ICR MS/MS and the overall procedure is lower time- and sample-consuming than traditional tandem mass spectrometry. In 2D MS more information is available at once through the observation of the neutral-loss lines and the vertical ion scans. The IRMPD 2D MS analysis of CaM resulted in a ~40% cleavage coverage of the protein under analysis. The cleavage coverage of CaM has been improved of more than 20% by performing a single additional 30-min BUP ECD 2D experiment, for a total of 68%.

In conclusion, 2D FT-ICR MS can be used for the identification of proteins in complex mixtures, although the technique could be improved for top-down analysis. The biggest limitation is currently represented by computational capacities.

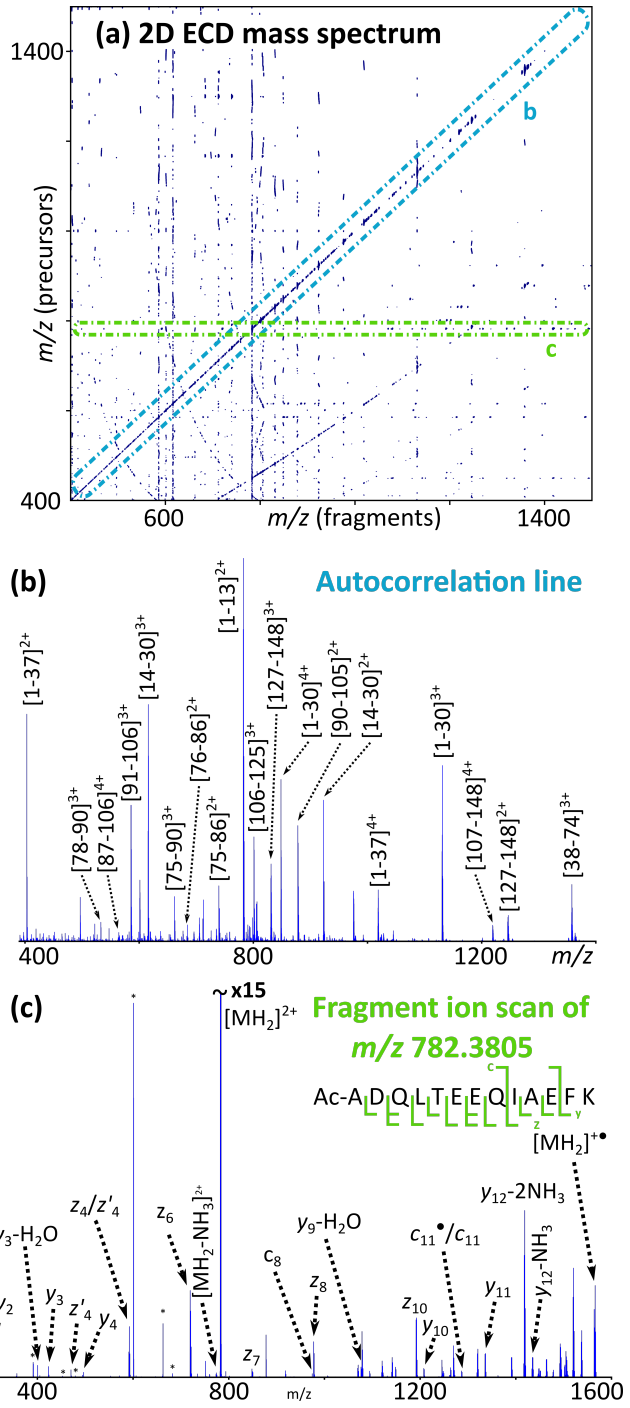



Figure 3.6.: Bottom-up ECD 2D MS analysis of calmodulin. (a) ECD 2D mass spectrum of tryptic digested CaM, with highlight on the autocorrelation line and a precursor ion scan. The vertical lines represent the residual scintillation noise, heavily reduced by urQRd. (b) Extraction of the autocorrelation line, with the assigned tryptic peptides. (c) Fragment ion scan of the peptide with $m/z$ 782.3805, with its sequence and relative cleavage coverage.

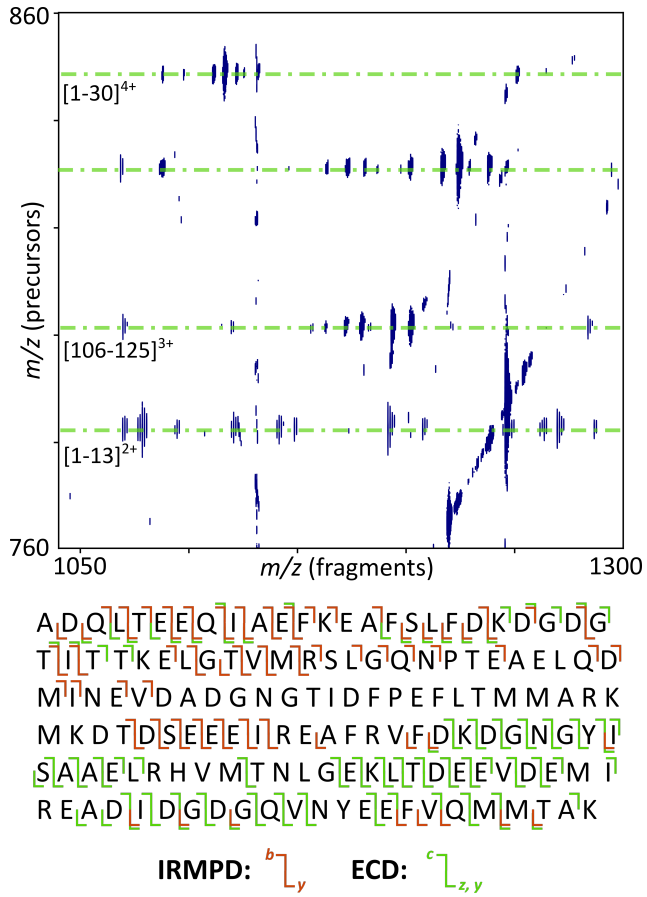



Figure 3.7.: Deep interpretation of a 2D mass spectrum. A zoom is reported (top), showing the horizontal lines of signals due to the fragmentation patterns of different peptides. (bottom) Cleavage coverage of CaM following 2D MS analysis. The total cleavage coverage reaches about 68%. The cleavages obtained with 2D ECD MS are reported in green (69/148 cleavages, ~47%). The cleavages in orange derive from the 2D IRMPD MS study of CaM (~40%).

# 4. Use of 2D MS for Top-down deep sequencing of proteins

The contents of this chapter have been submitted for publication in *Analytical Chemistry* (method development on Ubiquitin) [190] and the *Journal of American Society of Mass Spectrometry* (application of MS/2D MS to Calmodulin) [191].

Two-dimensional Fourier Transform Ion Cyclotron Resonance Mass Spectrometry (2D FT-ICR MS or 2D MS) allows data independent fragmentation of all ions in a sample and correlation of fragment ions to their precursors through the modulation of precursor ion cyclotron radii prior to fragmentation. Developments in computer capabilities and implementations in FTMS over of the last decade have allowed 2D FT-ICR MS to become a useful analytical tool for the examination of macromolecules in bottom-up proteomics, and more recently in top-down protein analysis. In this work, a new method of top-down protein analysis is developed using two-dimensional mass spectrometry, called MS/2D FT-ICR MS or MS/2D MS. MS/2D MS is an experiment equivalent to $MS^3$ in which quadrupole ion selection is used to select a precursor charge state for a protein; the ions are then fragmented using CAD in the hexapole collision cell, and subsequently 2D MS is used to fragment all precursor ions using IRMPD and ECD in the ICR cell. MS/2D MS is used in this work to achieve extensive inter-residue bond cleavage and assignment for ubiquitin and calmodulin, noting several unique features for fragment identification.

## Significance

The complexity of protein forms determined the rise of Proteomics as a field of study above Genomics, and technologies such as mass spectrometry (MS) offer a leading platform for characterisation of such macromolecules. Studying proteins through MS departing from their entirety (Top-Down Proteomics, TDP), rather than from smaller peptides obtained through their enzymatic digestion (Bottom-Up Proteomics, BUP), offers the highest amount of structural information, but constitutes a more challenging experiment often leading to a low sequence coverage of the protein in analysis. In this work, a new method for TDP is developed using two-dimensional mass spectrometry, called MS/2D MS. MS/2D MS aims at achieving extensive sequencing of proteins in TDP, delivering highest possible amount of information in a single experiment.

## 4.1. Introduction

The increasing interest in the discovery of the various proteoforms that lead to divergent protein functionalities in different biological systems raised the importance of a more detailed characterisation of the macromolecules that constitute the study of proteomics [211, 212]. Although the complexity of proteoforms has been showed with different techniques [213], technologies such as mass spectrometry are able to achieve protein characterisation at a sequence level, highlighting the differences due to post-translational modifications (PTM's) [41]. Amongst the mass spectrometry techniques, Fourier transform ion cyclotron resonance mass spectrometry (FT-ICR MS) has proven to be an essential analytical platform for proteomics, offering the highest resolving power and mass accuracy for the determination of the mass-to-charge ratios ($m/z$) of both intact and digested molecules [3, 6]. MS methodologies offer in fact two main approaches for the analysis of proteins, called respectively "top-down" and "bottom-up" [214, 82].

As explained in Chapter 1.6, the top-down approach offers the most copious data for the identification of a protein, whose important characteristics such as structural information, site-specific mutations, and PTM's are not affected by the extreme sample-preparation conditions to which the analytes are subjected within the bottom-up approach. However, the more complex data generated and technical limitations currently limit this method's execution.

FT-ICR MS alleviates technical limitations tied to the high resolving power required for unambiguous characterisation of small mass shifts due to PTM's such as deamidation ($\Delta$m $\approx$ 1 Da on a 15-kDa protein) and to resolve overlapping fragment peaks, and offers a high mass accuracy for high-confidence analysis of MS/MS spectra [215, 6, 214, 162]. Furthermore, with its two-dimensional development, 2D

FT-ICR MS or more simply 2D MS, it has been able to face the challenges imposed to every tandem mass spectrometry technique by isolation of precursor ions before the mass analyser. As mentioned, in fact, to obtain the highest amount of structural information, proteins need to be fragmented through tandem mass spectrometry. Tandem MS goes often beyond the MS/MS strategy, involving the further fragmentation of already generated fragments ($MS^n$). In order to generate interpretable tandem mass spectra, the activation stage necessitates a high quality isolation of the precursor. This circumstance is due to the impossibility to determine the origin of the fragments if in the spectrum there is more than one precursor, and, depending on the complexity of the mixture, it is not always possible to achieve. In a top-down MS/MS analysis with relatively pure samples, such isolations are relatively easy, but they turn into a problem for further analysis by MS/MS/MS or $MS^3$, due to the complexity of the generated primary fragmentations. Furthermore, it stands to reason that isolating every single peak in the MS/MS spectrum to generate a series of MS/MS/MS spectra can be sample- and time-consuming. Two-dimensional FT-ICR MS removes the need for isolations in MS/MS analysis as a result of its ability to correlate precursor and fragment ions based on the modulation of their fragmentation in the ICR-cell [172, 216, 217, 218]. In 2D MS the fragmentation patterns of all the molecules in a mixture are obtained in a single experiment.

Principles of 2D MS have been explained in previous studies and can be found in Chapter 1.10.

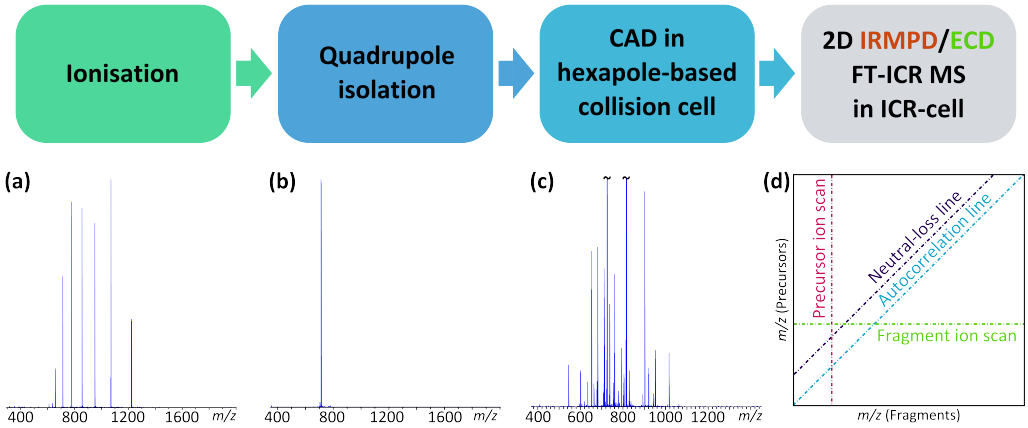



Figure 4.1.: Workflow of MS/2D MS with representative spectra from the different steps of the procedure. Proteins are protonated in the ionisation source, generating a protein charge-state distribution like the one in insert (**a**). A charge state of interest is then isolated in the quadrupole (**b**) and fragmented with CAD in the hexapole collision cell, generating primary fragment ions shown in (**c**). The generated CAD-fragments are finally sent to the ICR-cell for two-dimensional MS analysis. (**d**) The autocorrelation line (blue) shows the correlation of the precursor ion signals with their own cyclotron radius. Horizontal fragment ion spectra (green) show the fragmentation patterns of each precursor ion. Vertical precursor ion spectra (magenta) show the precursor ions of each fragment ion. Neutral-loss lines (purple) show the loss of neutrals by precursor ions.

Since 2D FT-ICR MS demonstrated the ability to decrease sample and time-consumption carrying sometimes more information than MS/MS proteomics studies [176, 165], it is here used to develop an MS/MS/MS-equivalent method for top-down detailed characterisation of proteins. The developed method is called MS/2D MS and combines the high-resolution 1D MS/MS fragmentation of FT-ICR MS with the ability to further characterise all the generated primary fragments at once using 2D FT-ICR MS. Figure 4.1 shows the workflow of an MS/2D MS experiment, with a representative mass spectrum for each step of the process (inserts a to d). In MS/2D MS, a charge state of interest of the protein under analysis (**a**) is firstly selected using quadrupolar isolation (**b**). The isolated ion species is then fragmented by acceleration into the collision cell using CAD (**c**), and the generated CAD-fragments are sent into the ICR-cell for two-dimensional mass spectrometry analysis using IRMPD or ECD as fragmentation techniques. A two-dimensional mass spectrum is generated (**d**), containing all the IRMPD/ECD fragmentation patterns of the CAD-fragments, constituting information equivalent to $MS^3$ experiments of each and every primary CAD-fragment ions. 2D MS is performed without any sort of in-cell isolation, and all the CAD-fragments' further activation with IRMPD or ECD is modulated as per principles of 2D FT-ICR MS. Other tandem mass spectrometry steps can be added

to the technique, as long as it is possible to further isolate and fragment an ion before it enters the ICR-cell. For example, for a relatively pure protein, in-source dissociation (ISD) could be used to dissociate all the charge states of the protein in analysis, followed by isolation, CAD MS/MS, and 2D FT-ICR MS, for a pseudo MS/MS/2D MS experiment.

MS/2D MS is here performed on modest size proteins such as Ubiquitin and Calmodulin. Ubiquitin (Ubi) is a relatively small protein (76 amino acids, 8.6 kDa) that has been extensively studied with FT-ICR MS using many different fragmentation techniques [142, 93, 219, 155]. It is often used as a model for proteomics because of its composition and the considerable knowledge about its behaviour in the gas phase as a multiply charged ion [220, 56, 221]. Calmodulin is a ubiquitous protein able of modulating several signalling processes in the human body by binding to up to four Calcium(II) atoms [201, 202]. As mentioned in the previous chapter, CaM has been used as a standard for deamidation [169] and to investigate the binding of anti-cancer drugs to proteins [209, 170]. Its dimensions (149 residues, 16.8 kDa) and relatively simple structure, given the absence of disulphide bonds, make it another good model for top-down proteomics [165].

This work represents an expansion on the multidimensional study of calmodulin started with the works reported in Chapter 3, with the goal of improving the structural characterisation of the macromolecule by using the unique features of two-dimensional mass spectrometry.

## 4.2. Experimental section

**Materials.** Calmodulin from bovine testes, ubiquitin from bovine erythrocytes, were purchased from Sigma-Aldrich (Dorset, UK). HPLC grade methanol and formic acid (HAc), were obtained from Fisher Scientific (Loughborough, UK). Water was purified by a Millipore Direct-Q purification system (Millipore, Nottingham, UK).

**Sample preparation.** Aqueous stock solutions of Calmodulin and Ubiquitin were prepared. Salt adducts were removed from each protein's aqueous solution through repeated centrifugation with 3 kDa molecular weight cut-off centrifugal filters (Amicon Ultra 0.5 mL, Merck Millipore Ltd, Tullagreen, Ireland) in a technico maxi microcentrifuge (Fisher Scientific, Loughborough, Leicestershire, UK), at 14000 rpm for 30 min. Aqueous solutions of CaM, and Ubi were then passed through 0.22 µm centrifugal filters to remove particulates prior to electrospray ionisation. Calmodulin (0.4 µM) was dissolved in a 75:25 water/methanol (v/v) solution with 0.3% (v/v) of formic acid. Ubiquitin (0.25 µM) was dissolved in a 50:50 water/acetonitrile (v/v) solution with 0.3% (v/v) of formic acid.

**Instrumentation.** FT-ICR MS was performed on a 12 T Bruker solariX FT-ICR mass spectrometer (Bruker Daltonik GmbH, Bremen, Germany). Samples were ionised using an electrospray ionisation (ESI) source at flow rates of 60-80 µL/h. When necessary, ions were accelerated in the source in order to remove eventual salt adducts – this is mainly the case of Calmodulin (see Table B.36 in Supporting information for the energy value used for this in-source dissociation), then they were transferred through two octopoles to a quadrupole, in which accumulation occurred. CAD was performed by accelerating the ions from the quadrupole to a hexapole collision cell with argon gas, then transferring the ions through a transfer hexapole (1.0-1.2 ms flight time) to an Infinity Cell [18], for mass analysis and detection.

**Control one-dimensional mass spectra acquisition.** One-dimensional MS spectra were acquired as controls for each sample, generating full (MS) and tandem (MS/MS and MS/MS/MS) mass spectra. The control spectra were acquired using 4M data-points (16 bits) for a transient length of 1.6777 s, over a mass range of $m/z$ 147.5-3000. MS control spectra were accumulated for 8 scans. MS/MS (or $MS^2$) control spectra were acquired isolating a specific charge state of the protein in analysis and fragmenting it through CAD. MS/MS spectra were acquired accumulating for 64 scans. Tandem MS/MS/MS (or $MS^3$) control spectra were generated using CAD as first fragmentation technique, followed by transfer of the generated fragments into the ICR-cell. In the cell, specific CAD-fragments were isolated using the Multi-CHEF capability of the SolariX system. This in-cell isolation was performed for each spectrum selecting a parent ion with isolation power at 69%, the range of ejection between $m/z$ 100 and 3000, a notch width of 4.0 s, and 0.158465 seconds-long shots at 5.0% of their power. The isolated fragments were then subjected to IRMPD or ECD to achieve further fragmentation. $MS^3$ spectra were acquired accumulating for 128 scans. IRMPD fragmentation was achieved using a continuous wave, 25 W, $CO_2$ laser (Synrad Inc., Washington, USA) held at 70% of its power output. IR photons were produced at a wavelength of 10.6 µm and pulsed into the ICR cell for 0.3-0.5 s prior to detection. ECD was performed generating electrons from a heated hollow cathode (1.5 A) and pulsating them at 10 V into the ICR-cell for 0.2 s prior to detection (ECD bias 1.5-2.0 V). See supporting information for conditions particular to each control spectrum (Tables B.32 and B.36).

**Two-dimensional mass spectra acquisition.** MS/2D MS spectra were acquired for each sample by isolation of a specific charge state of the protein in analysis, fragmentation of such ion using hexapole-based CAD, and transfer of the generated fragments to the ICR-cell for further fragmentation and analysis through two-dimensional mass spectrometry. For each sample, CAD fragments were generated isolating the ion of interest in the instrument's quadrupole with an isolation window of 20 $m/z$.

The ions of interest were respectively *m/z* 714.308671 ($MH_{12}^{12+}$) for Ubi, and *m/z* 1199.568090 ($MH_{14}^{14+}$) for CaM. Isolated ions were then accelerated at 12-16 V into Argon gas present in the hexapole-based collision cell, accumulated for 1.0-2.0 s, and sent into the ICR-cell. Two-dimensional mass spectrometry was performed on the generated CAD-fragments using IRMPD and ECD as fragmentation techniques, modulating their fragmentation according to the scheme in Figure S1 in the Supporting Information. $P_1$ and $P_2$, as per principles of 2D FT-ICR MS, are two identical pulses. They were held at 15% of excitation power (70 $V_{pp}$ amplitude), and their pulse length was optimised according to previous studies [178]. For the experiment with IRMPD, $P_1$ and $P_2$ were set at a length of 0.4 µs. For the ECD experiment, their pulse length was set at 4.0 µs. The incremental time $t_1$ was set at an initial delay depending on the frequency correspondent to the lowest *m/z* analysed in the vertical dimension, and it was incremented by the same amount at each reiteration of the pulse programme. For Ubiquitin, for example, the initial delay $t_1$ was 1.0 µs, incremented 2048 times by 1.0 µs. For the CAD/IRMPD MS/2DMS analysis of Ubi, 2048 scans (increments of $t_1$) of 512k data points were acquired over a mass range of *m/z* 368.2-3000 on the vertical axis and *m/z* 147.5-3000 on the horizontal axis, the total time of acquisition was ~80 minutes. For the CAD/ECD MS/2DMS analysis of Ubi, 2048 scans of 512k data-points (16 bits) were acquired over a mass range of *m/z* 368.2-3000 on the vertical axis and *m/z* 147.5-3000 on the horizontal axis; total time of acquisition was ~80 min. For the CAD/IRMPD MS/2DMS analysis of CaM, 2048 scans of 512k data-points were acquired over a mass range of *m/z* 368.2-3000 on the vertical axis and *m/z* 147.5-3000 on the horizontal axis, the total time of acquisition was ~80 minutes. For the CAD/ECD MS/2DMS analysis of CaM, 1024 scans of 2M data points were acquired over a mass range of *m/z* 368.2-3000 on the vertical axis and *m/z* 147.5-3000 on the horizontal axis; total time of acquisition was ~80 min.

**Data processing and analysis.** The one-dimensional control spectra were externally calibrated with the solariX control software (Bruker Daltonics, Billerica, MA, USA) using Agilent ESI-L Low Concentration Tuning Mix (Agilent Technologies, Stockport, UK), then internally calibrated using theoretical *m/z* of known fragment peaks with a quadratic calibration function in the Bruker Data Analysis v4.1 software (Bruker Daltonics GmbH, Bremen, Germany).

The 2D mass spectra were processed using SPIKE [192] and the urQRd algorithm for denoising [183]. For the generated spectra, denoising ranks between 20 and 30 were used. The spectra were externally calibrated in the vertical and horizontal dimensions using the same parameters set for the respective one-dimensional control spectra. Internal calibration was performed in the vertical dimension through the extraction of the autocorrelation line and the use of theoretical mass-to-charges of known CAD-fragments, and in the horizontal dimension using theoretical *m/z* of

known IRMPD or ECD fragments. One-dimensional and two-dimensional spectra were manually interpreted and assigned. The full list of parameters used for the two-dimensional mass spectrometric analysis of Ubi and CaM and their one-dimensional homologous control spectra are provided in the Supporting Information (Tables B.32 and B.36).

## 4.3. Results

### 4.3.1. Ubiquitin

Figure 4.1.a shows the full ESI/FT-ICR mass spectrum of ubiquitin (Ubi) in denaturing conditions acquired as control. The spectrum was acquired in 4M data-points (16 bit) for a transient of 1.6777 s, and externally calibrated before acquisition with standard mixture. The derived magnitude mode resolving power is ~340000 at $m/z$ 400. Figure 4.1.a shows the range of charge states reached by Ubi in denaturing conditions: the charge state corresponding to $m/z$ 714.3087 ($MH_{12}^{12+}$), chosen for the high relative abundance and number of charges, was isolated and accumulated (Figure 4.1.b) for subsequent fragmentation with CAD, generating the spectrum in Figure 4.1.c, elucidated below.

#### 4.3.1.1. CAD MS/MS and CAD/(IRMPD/ECD) $MS^3$ of ubiquitin

Figure 4.1.c shows the 1D CAD tandem mass spectrum of ubiquitin acquired for comparison with the MS/2D MS analysis. The spectrum was acquired in 4M data-points (16 bit) for a transient of 1.6777 s, and calibrated internally using theoretical $m/z$ of known CAD fragments. The resulting resolving power is ~440000 at $m/z$ 400. Typical CAD fragments could be assigned with a high average Mass Measurement Accuracy of 0.19±1.53 ppm (MMA±SD). The vast majority of the assigned fragments consisted in $b/y$ ions, but also some internal fragments could be recognised. All the successfully assigned peaks for the CAD MS/MS spectrum of Ubi are listed in Table B.1 in the Supporting Information. The resulting cleavage coverage is ~33.4%.

As a control for the MS/2D MS analysis, a CAD fragment ion, $y_{58}^{9+}$ ($m/z$ 726.2838), was isolated in the ICR-cell, and further fragmented with IRMPD and ECD. The resulting MS/MS/MS spectra were acquired in 4M data-points (16 bit) for a transient of 1.6777 s, and calibrated internally using theoretical $m/z$ of the expected known IRMPD or ECD fragments. The reached resolving power for the spectra is ~390000, reported at $m/z$ 400. All the present fragments could be assigned with a high MMA of 0.08±0.21 ppm for the CAD/IRMPD $MS^3$ spectrum, and of 1.42±1.39 ppm for the CAD/ECD $MS^3$ spectrum. The correctly assigned fragments are listed in Table B.2

for the CAD/IRMPD spectrum and in Table B.3 for the CAD/ECD $MS^3$ spectrum in the Supporting Information. The CAD/IRMPD and CAD/ECD $MS^3$ spectra of Ubi in denaturing condition with labelled assigned peaks and the respective cleavage coverage maps are shown in Figure 4.2.

#### 4.3.1.2. CAD/IRMPD MS/2D MS of ubiquitin

Figure 4.3 shows the two-dimensional mass spectrum resulting from the CAD/IRMPD MS/2D MS analysis of ubiquitin in denaturing conditions (**a**), with the extraction of the autocorrelation line (**b**) and a fragment mass spectrum (**c**). The 2D mass spectrum of Figure 4.3.a was acquired with 2k data-points (2048 iterations in $t_1$, scans) in the vertical dimension and 512k data-points (16 bit) in the horizontal dimension. The vertical resolving power, $RP_v$ is ~600, reported at $m/z$ 400. The horizontal resolving power $RP_h$ is ~170000 at $m/z$ 400. The total resolving power of the mass spectrum is $RP_{2D} \approx 100$ M, reported at the same $m/z$. Signals were calibrated before the acquisition using an external calibration with a standard mixture and a quadratic calibration. The mass spectrum shows a clear autocorrelation line and several other peaks, resulting from the fragmentation of the precursor ions with IRMPD. Some other lines of signal represent harmonics, or noise frequencies captured during the acquisition phase. The elimination of the scintillation noise [180] using of the urQRd algorithm in the processing of the two-dimensional mass spectrum allows a clear visualisation and interpretation.

The autocorrelation line is extracted and shown in Figure 4.3.b. As expected, the extraction of the autocorrelation line generated a mass spectrum of the precursors, in this case representing the primary fragments generated by activation in the hexapole collision cell with CAD. The autocorrelation line thus shows typical $b/y$ ions, some $a$ ions, and some internal ions generated by such fragmentation technique, as well as several neutral losses (water or ammonia). The spectrum is very dense, as expected for top-down MS/MS spectra, but the resolution reached with the MS/2D MS experiment is enough to isotopically resolve the fragment ion peaks, and allow manual assignment. The autocorrelation line was calibrated with theoretical mass-to-charge ratios of known CAD fragments, and the peaks could be assigned with an average mass accuracy of 0.39±3.85 ppm. In Figure 4.3.b, the percentage of $b/y$ ions, internal ions, neutral losses assigned, and unknown ions found is reported for the autocorrelation line. It is hypothesised that unassigned ions occur from secondary fragmentation or even from "sequence scrambled" ions, as common for CAD or IRMPD MS/MS techniques [89]. The cleavage coverage equivalent to the top-down MS/MS analysis of Ubi can be determined by calculating the cleavage coverage of the autocorrelation line. The resulting cleavage coverage, counting the cleavages assigned to internal ions, is ~45.3%. Cleavages corresponding to the reported cleavage coverage are represented in blue on the cleavage coverage map in Figure 4.4.d. A list of the complete

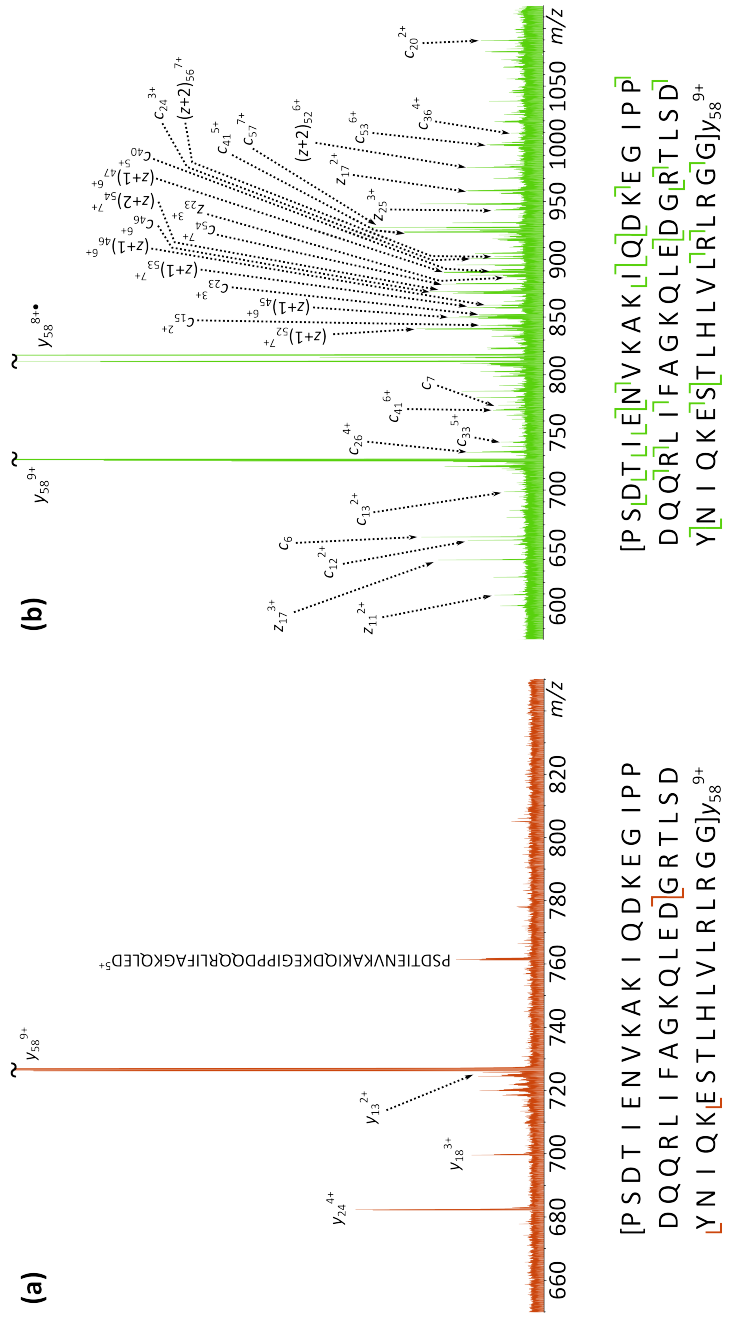



Figure 4.2.: CAD/IRMPD (a) and CAD/ECD (b) MS/MS/MS spectra of Ubiquitin, obtained fragmenting the molecule with CAD in the hexapole-based collision cell, and isolating the ion $y_{58}{}^{9+}$ for further fragmentation in the ICR-cell respectively with IRMPD or ECD. Cleavage coverage maps for the fragmented ions are reported under the respective spectra.

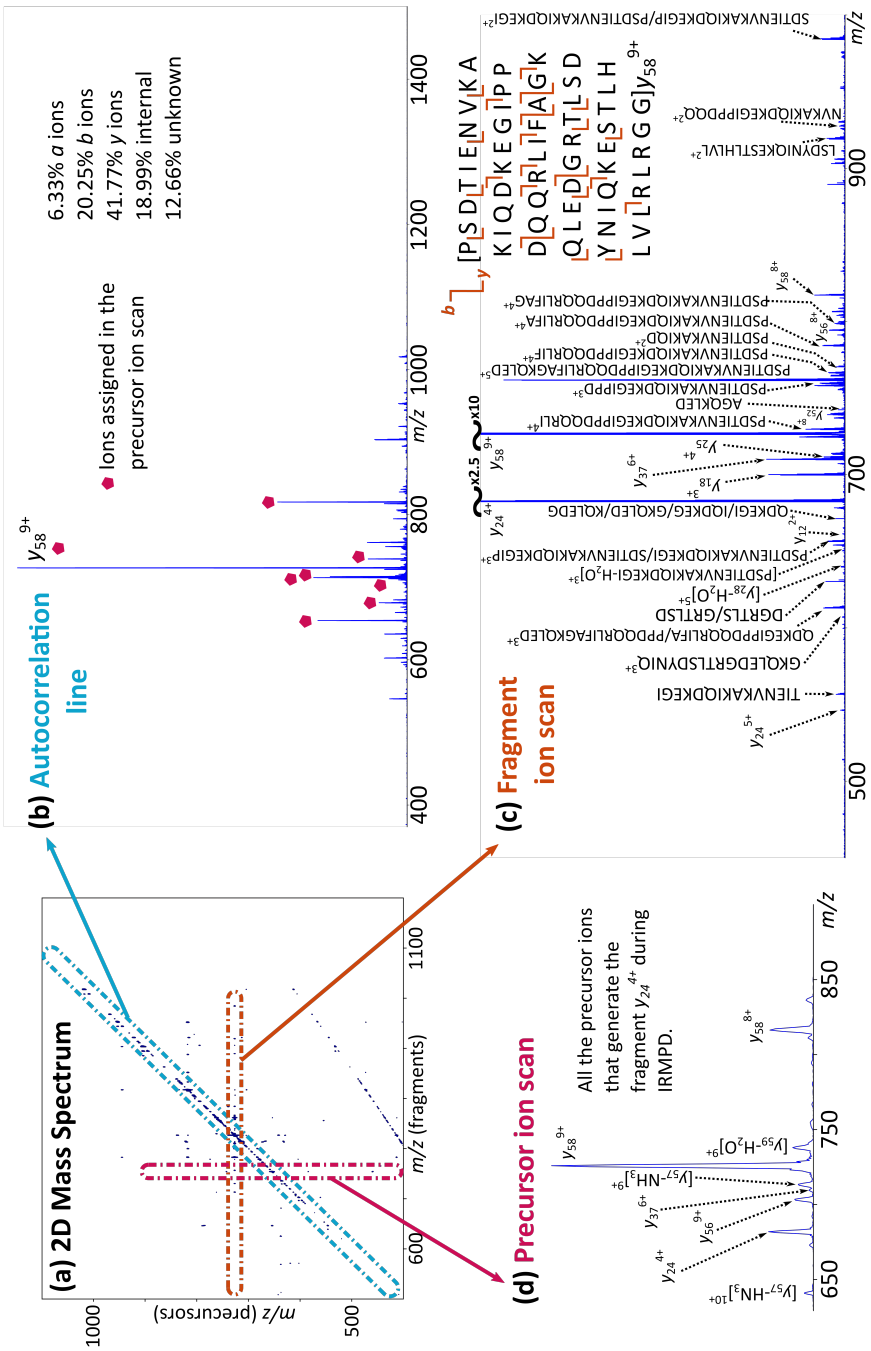



Figure 4.3.: CAD/IRMPD MS/2D MS of Ubiquitin in denaturing conditions. (**a**) Two-dimensional mass spectrum, output of the CAD/IRMPD MS/2D MS analysis of Ubi. The autocorrelation line and horizontal fragmentation patterns are observed. The curved line departing from the bottom of the spectrum constitutes harmonics signal. (**b**) Extraction of the autocorrelation line, with the percentages of the correctly assigned ions. Labelled ions represent the ions assigned through the precursor ion scan showed in spectrum d; they include ion $y_{58}{}^{9+}$, whose fragmentation pattern is shown in spectrum c. (**c**) Fragment ion scan of the ion $y_{58}{}^{9+}$ ($m/z$ 726.2838) and respective cleavage coverage. Non-diagnostic neutral losses are not labelled for better visualisation. (**d**) Precursor ion scan of the ion $y_{24}{}^{4+}$ ($m/z$ 682.3820), constituting important information for the assignment of y ions on the autocorrelation line (**b**).

peak assignments is available in Table B.4 in the Supporting Information. The dynamic range observed for the autocorrelation line is ~3300.

Figure 4.3.c shows the extraction of a fragment ion scan, corresponding to the fragmentation pattern of the ion ${y_{58}}^{9+}$, *m/z* 726.2838. The generated mass spectrum shows all the fragments of the ion ${y_{58}}^{9+}$ generated by IRMPD inside the ICR-cell, and shows $b/y$ ions, internal ions, and neutral losses of water and ammonia typical of infrared multiphoton dissociation. The horizontal spectrum was internally calibrated using the theoretical *m/z* of the known fragments, and the same calibration has been used for the entire two-dimensional mass spectrum for the assignments of the other extracted horizontal lines. The assigned peaks are labelled in Figure 4.3.c, and the cleavage coverage is reported and equal to ~40.3% for the fragmented ion. Horizontal fragment ion scans were extracted for all the successfully assigned precursors, giving a cleavage coverage of ~62% for the CAD/IRMPD MS/2D MS experiment of ubiquitin at the specified conditions. A cleavage coverage map for the experiment is reported in Figure 4.4.d. A complete list of the precursors that reported significant fragmentation is included in the Supporting Information, along with the complete peak assignments for their fragmentation patterns (Tables B.5 to B.14).

Figure 4.3.d shows an example of a mass spectrum generated by the extraction of a vertical precursor ion scan. Precursor ion scans are embedded in 2D MS data and allow simultaneous detection of all the precursor ions of a selected fragment in order to better understand the involved fragmentation mechanisms. In this case, all the precursor ions that generate the ion ${y_{24}}^{4+}$ are shown and labelled in the mass spectrum indicating that all these precursor ions must be C-terminal ions higher than $y_{24}$. Precursor ions were assigned through cross-correlation with the autocorrelation line. The $y$ ions assigned through the extraction of the vertical ion scan are labelled in Figure 4.3.b with magenta pentagons.

#### 4.3.1.3. CAD/ECD MS/2D MS of ubiquitin

Figure 4.4 shows the two-dimensional mass spectrum resulting from the CAD/ECD MS/2D MS analysis of ubiquitin in denaturing conditions (**a**), with the extraction of the autocorrelation line (**b**) and a fragment mass spectrum (**c**). The 2D mass spectrum of Figure 4.4.a was acquired with 2k data-points in the vertical dimension and 1M data-points (16 bit) in the horizontal dimension. The vertical resolving power, $RP_v$ is ~600, reported at *m/z* 400. The horizontal resolving power $RP_h$ is ~190000 at *m/z* 400. The total resolving power of the mass spectrum is $RP_2D \approx$ 105 M, reported at the same *m/z*. As for the CAD/IRMPD MS/2D MS experiment, signal acquisition was calibrated externally using a standard mixture and a quadratic calibration. The mass spectrum shows a clear autocorrelation line and several other peaks, in this case resulting from the fragmentation of the precursor ions with ECD.

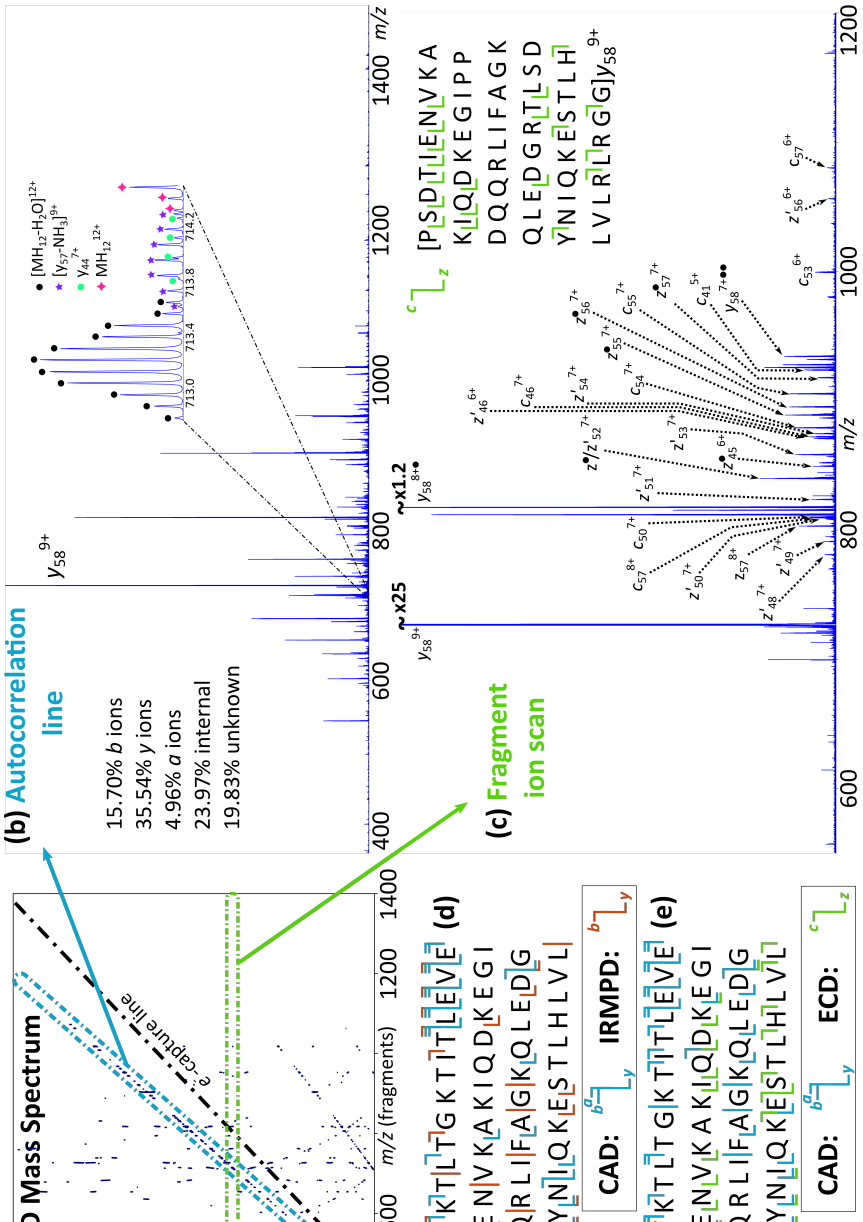



Figure 4.4.: CAD/ECD MS/2D MS of Ubiquitin in denaturing conditions. (**a**) Two-dimensional mass spectrum, output of the CAD/ECD MS/2D MS analysis of Ubi. The autocorrelation line and horizontal fragmentation patterns are observed. An electron-capture line is highlighted. The curved line at the bottom of the spectrum is a harmonic. (**b**) Extraction of the autocorrelation line, with the percentages of the assigned precursor ion types. The insert shows the complexity of the spectrum, typical of a top-down MS/MS experiment. Ion $y_{58}{}^{9+}$, whose fragmentation pattern is shown in spectrum c, is labelled. (**c**) Fragment ion scan of the ion $y_{58}{}^{9+}$ ($m/z$ 726.2838) and respective assigned cleavages. Non-diagnostic neutral losses are not labelled for better visualisation. A great number of $c/z$-type ions is observed. (**d**) Cleavage coverage map of the CAD/IRMPD MS/2D MS analysis of Ubi. (**e**) Cleavage coverage map of the CAD/ECD MS/2D MS analysis of Ubi. For both the reported maps, blue labels represent the fragments obtained by CAD and assigned analysing the autocorrelation line. Orange labels (d) indicate IRMPD fragments and green labels (e) indicate ECD fragments assigned through the analysis of horizontal scans.

An electron-capture line can be highlighted on the spectrum, showing the signals of all the precursors that captured an electron during the activation period in the ICR-cell without undergoing subsequent fragmentation. All the peaks in the electron-capture line correspond to radical cation charge-reduced precursors. Other lines of signal representing harmonics or phantom frequencies captured during the acquisition phase can also be observed. In this case, although minimised through the use of the urQRd denoising, some vertical lines corresponding to residual scintillation noise can be observed, but their recognition does not hinder spectral assignment of real peaks.

The autocorrelation line was again extracted, and it is showed in Figure 4.4.b. As expected, the autocorrelation line shows typical fragments generated by collisionally-activated dissociation: $b/y$ ions, some a ions, internal ions, and several neutral losses of water or ammonia. Although acquired in the exact same conditions of ionisation, ion transfer, and collisional energy, direct comparison between the autocorrelation lines extracted for the CAD/IRMPD (Figure 4.3.b) and the CAD/ECD MS/2D MS experiments (Figure 4.4.b) show visual differences in assigned ions and their signal intensity. In 2D FT-ICR MS, in fact, signal intensities of peaks on the autocorrelation line are dependent not only by the relative ion abundance, but also by the variation of intensity of the precursor during the fragmentation period. The spectrum in Figure 4.4.b is again very dense as expected for top-down MS/MS spectra, but the resolution reached with the MS/2D MS experiment is enough to isotopically resolve the fragment ion peaks. The autocorrelation line was calibrated with theoretical mass-to-charge ratios of known CAD fragments, and the peaks could be assigned with an average mass accuracy of 0.09±2.55 ppm. Figure 4.4.b reports the percentage of $b/y$ ions, internal ions, neutral losses assigned, and unknown ions found for the autocorrelation line. The resulting cleavage coverage from the autocorrelation line, counting the cleavages addressed to internal ions, is ~45%, similar to Figure 4.3.b. Cleavages corresponding to the reported cleavage coverage are represented on the cleavage coverage map in Figure 4.4.e, in blue. A list of the complete peak assignments is available in Table B.15 in Supporting Information. The dynamic range observed for the autocorrelation line is ~3700.

To allow direct comparison between MS/2D MS using different activation techniques, the fragmentation pattern of the ion ${y_{58}}^{9+}$ ($m/z$ 726.2838) is extracted from the 2D mass spectrum through a fragment ion scan, and showed in Figure 4.4.c. Similar to the previous case, the generated mass spectrum shows all the fragments of the ion ${y_{58}}^{9+}$ generated during the fragmentation period inside the ICR-cell, in this case through the use of ECD as the fragmentation technique. It is in fact possible to recognise $z^{\bullet}/z$' ions, $y$ ions, charge-reduced species, neutral losses of water and ammonia, and $c$ ions, whose masses were calculated through the structure proposed in the scheme of Figure B.1 in the Supporting information. The mentioned fragments are typical of ECD fragmentations, and constitute complementary information to

IRMPD. The horizontal spectrum was internally calibrated using theoretical mass-to-charge ratios of the known $z$ and $y$ fragments, and the same calibration has been used towards the entire two-dimensional mass spectrum for the assignments of the other extracted horizontal lines. The assigned peaks are labelled in Figure 4.4.c, and the cleavage coverage is reported and equal to ~28.1% for the ${y_{58}}^{9+}$ ion. Horizontal fragment ion scans were extracted for all the successfully assigned precursors, giving a total cleavage coverage of ~72% for the CAD/ECD MS/2D MS experiment of ubiquitin at the specified conditions. A cleavage coverage map for ion the experiment is reported in Figure 4.4.e. A complete list of the precursors that reported significant fragmentation is included in the Supporting Information, along with the complete peak assignments for their fragmentation patterns (Tables B.16 to B.19).

Combining the data obtained with the CAD/IRMPD and CAD/ECD MS/2D MS experiments, a cumulative cleavage coverage of 83% has been obtained for ubiquitin at the detailed conditions.

## 4.3.2. Calmodulin

Similarly to the analysis of Ubiquitin, one-dimensional control spectra were acquired for the full MS and CAD MS/MS analysis of Calmodulin.

### 4.3.2.1. Full MS of CaM

A one-dimensional full mass spectrum of Calmodulin (CaM, 17kDa) in denaturing conditions was acquired as control. Equivalently to the analysis of Ubiquitin, the spectrum was acquired using 4M data-points (16 bits), for a transient length of 1.6777 s, and a resulting resolving power of ~340000. Calibration was achieved through instrument calibration with a quadratic curve, using an external calibrant before the data acquisition. The full mass spectrum of CaM in denaturing conditions is shown in Figure 4.5. It exhibits a wide range of charge states reached by calmodulin at the detailed conditions, from 7+ to 18+. The most abundant charge state, corresponding to ${MH_{14}}^{14+}$ ($m/z$ 1199.5681), was isolated and accumulated in the quadrupole, for MS/MS fragmentation with CAD in the hexapole collision cell.

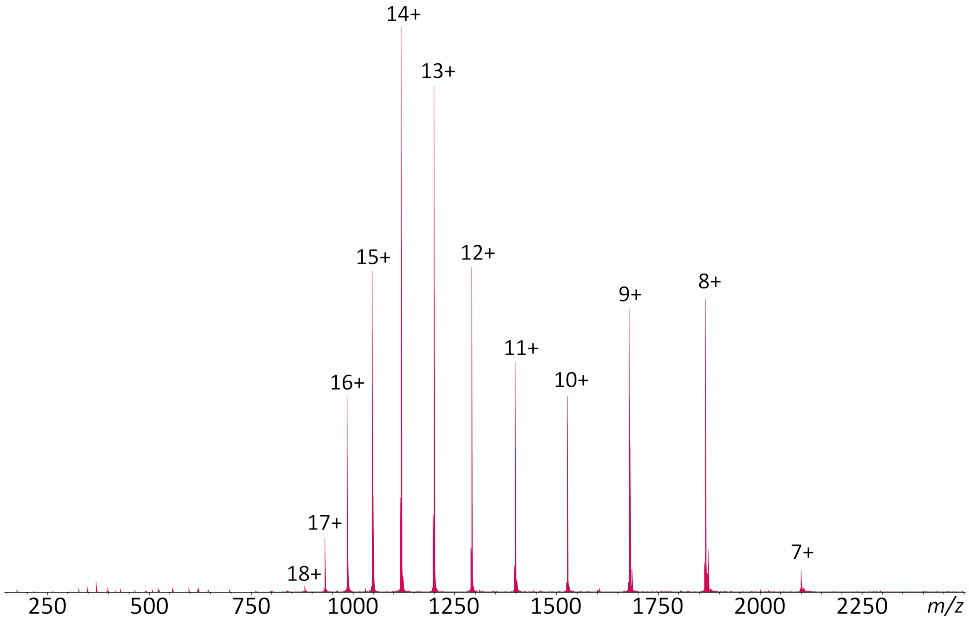



Figure 4.5.: Full mass spectrum of calmodulin, showing the range of charge states reached by the protein in denaturing conditions. Transients have been accumulated for 8 scans.

#### 4.3.2.2. CAD MS/MS of CaM

A one-dimensional CAD MS/MS spectrum for CaM was acquired with the same acquisition parameters listed for the full MS (Figure 4.6.a). The resulting resolving power was ~340000. The mass spectrum was internally calibrated using theoretical mass-to-charge ratios of the recognised peaks. CAD of calmodulin exhibited a large number of multiply charged $b/y$ ions (the vast majority represented by $b$ ions) matched by neutral losses in large scale. It is hypothesised that the unassigned peaks are originated by "sequence scramble", typical of $b$ ions in MS/MS experiments involving techniques such as CAD and IRMPD. Ions could be assigned with a high average mass accuracy of 0.21±0.98 ppm, leading to a cleavage coverage of 37.8% considering only $b$ and $y$ ions, and 41.2% including the assigned internal ions. A complete list of the assigned fragments for the 1D CAD MS/MS spectrum of CaM in denaturing conditions is reported in the Table B.33. A cleavage coverage map is reported in Figure 4.6.b.

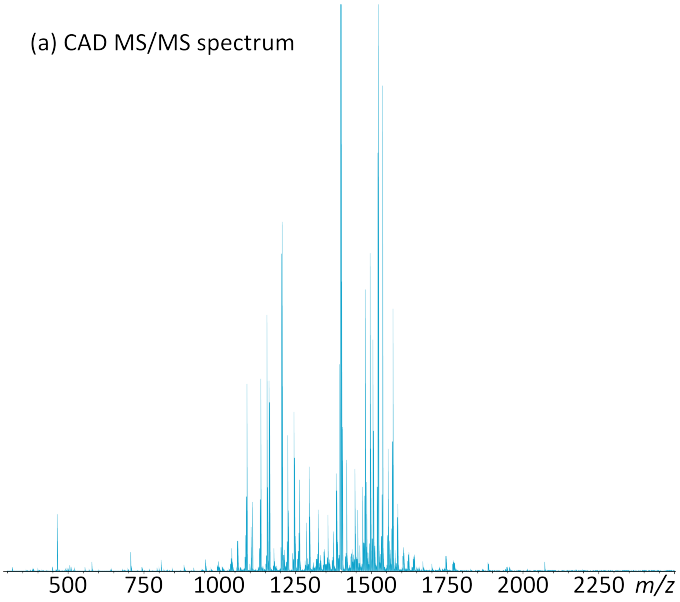



Figure 4.6.: (a) CAD MS/MS spectrum of calmodulin. It has been obtained isolating in the quadrupole of the mass spectrometer the 14+ charge state of calmodulin and accelerating the ions into Argon in a hexapole-based collision cell. Mass analysis and detection were performed in the ICR-cell. (b) Cleavage coverage map of the spectrum reported in (a): the cleavage coverage reaches 37.8% considering the generated b/y ions, and raises to 42.1% if internal ions are added. Complete mass assignments for the showed spectrum and coverage map are shown in table S.22.

### 4.3.2.3. CAD/IRMPD MS/2D MS of calmodulin

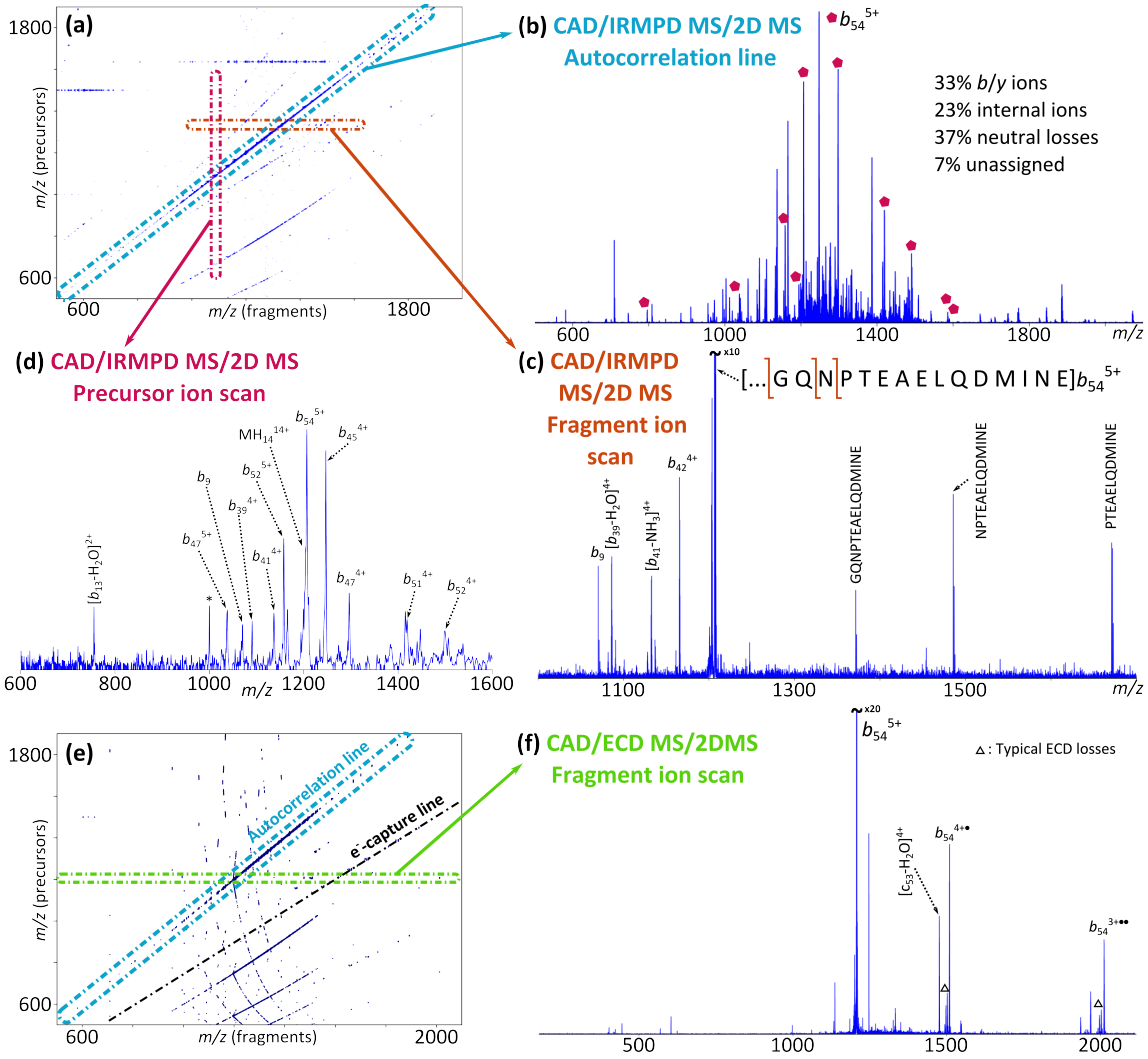



Figure 4.7.: CAD/ECD and CAD/IRMPD MS/2D MS of Calmodulin in denaturing conditions. (**a**) Two-dimensional mass spectrum for the CAD/IRMPD MS/2D MS analysis of CaM. The autocorrelation line, a horizontal scan and a vertical scan are highlighted and extracted respectively in spectra b, c, and d. (**b**) Autocorrelation line. (**c**) Fragment ion scan of the CAD-ion $b_{54}{}^{5+}$. (**d**) Vertical ion scan for the ion $b_9$. The assigned ions are labelled on the autocorrelation line (spectrum b) with a pentagon. (e) 2D mass spectrum from the analysis of CaM with CAD/ECD MS/2D MS. A horizontal scan is extracted in spectrum f. (f) Fragment ion scan of the ion $b_{54}{}^{5+}$.

Figure 4.7.a shows the two-dimensional mass spectrum resulting from the CAD/IRMPD MS/2D MS analysis of CaM, with the extraction of the autocorrelation line (**b**), a fragment ion scan (**c**), and a precursor ion scan (**d**). The characteristics of this spectrum are reported in the Supporting Information.

The autocorrelation line is extracted and shown in Figure 4.7.b. The spectrum generated with the extraction of the autocorrelation line shows the typical complexity of acquired top-down spectra, but even in this case the resolution is sufficient to isotopically resolve the present peaks. Signals on the autocorrelation line were calibrated internally, using the theoretical $m/z$ of the known fragments. The precursors on the autocorrelation line could be assigned with an average mass accuracy of 0.10±3.11 ppm. The autocorrelation line showed a large number of $b$ and $y$ ions, neutral losses, and internal fragments: typical ion fragments generated by techniques such as CAD. Figure 4.7.b indicates the percentages of the found ions' types. A complete list of the ions assigned on the autocorrelation line is reported in the Supporting Information (Table B.34).

Figure 4.7.c shows the extraction of a horizontal ion scan, corresponding to the fragmentation pattern of the CAD fragment ${b_{54}}^{5+}$ ($m/z$ 1206.3786), generated during the fragmentation period of the 2D analysis, inside the ICR-cell. Many $b$ and internal ions and neutral losses could be recognised in the spectrum, as expected for a fragmentation obtained by infrared multiphoton dissociation. The horizontal scan of ion ${b_{54}}^{5+}$ was calibrated using the theoretical $m/z$ of the recognised fragments, and the obtained fitting parameters were used to calibrate the entire two-dimensional mass spectrum with a quadratic equation. Fragment ions could be assigned with a mass accuracy of 0.21±0.98 ppm. The main $b$ and internal ions are labelled in Figure 4.7.c. Neutral losses of water and ammonia are present and can be observed, unlabelled. Figure 4.7.c reports the cleavage coverage calculated considering the labelled ions in the figure, corresponding to a value of 7.5% for ion ${b_{54}}^{5+}$. Horizontal fragment ion scans were extracted for all the successfully assigned precursors, and their fragmentation patterns assigned and used to calculate the cleavage coverage of CaM with the CAD/IRMPD MS/2D MS experiment. The total cleavage coverage is ~41%, calculated considering also the cleavages due to internal fragmentation, and showed in the cleavage coverage map of Figure 4.8.a.

As another example of capability of 2D FT-ICR MS techniques, a vertical scan is extracted and showed in Figure 4.7.d. This mass spectrum shows all the precursor ions that generate the ion $b_9$ ($m/z$ 1070.5000) during the fragmentation period in the ICR-cell. Figure 4.7.d shows about 15 precursors, which can only be N-terminal precursor ions and the molecular ion ${MH_{14}}^{14+}$ of CaM. Precursor ions were assigned through cross-correlation with the autocorrelation line. Signals due to experimental noise are labelled with a star symbol (*). The concurrent detection of all precursor ions and its ready availability is a powerful tool of 2D FT-ICR MS, and it is here used for simultaneous discrimination between N-terminus and C-terminus fragment ions. The $b$ ions assigned through the extraction of the vertical ion scan are labelled in Figure 4.7.b with a pentagon.

#### 4.3.2.4. CAD/ECD MS/2D MS of calmodulin

Figure 4.7.e shows the two-dimensional mass spectrum resulting from the CAD/ECD MS/2D MS analysis of Calmodulin in denaturing conditions.

The autocorrelation line was extracted and calibrated internally using theoretical $m/z$ of the recognised fragments. Several $b/y$ ions, internal fragment ions and neutral losses could be assigned, with an average mass accuracy of 2.42±4.66 ppm. Percentages of the assigned ions are reported in Figure B.3 in the Supporting Information, together with a complete list of the correct peak assignments for the autocorrelation line of the CAD/ECD MS/2D MS of CaM in denaturing conditions (Table B.35). Although the autocorrelation lines were obtained with the same CAD and ion transfer parameters for the two experiments they are mostly similar for the assigned fragments, but they differ in signal intensity, and some different precursors can be found in each of them (see tables B.34 and B.35). The cleavage coverage obtained for the CAD fragmentation of CaM, calculated on the autocorrelation line considering cleavages due to internal fragmentation, is of ~31%.

Figure 4.7.f shows the extraction of a fragment ion scan. Ion ${b_{54}}^{5+}$ was chosen as example to highlight the different fragmentation patterns obtained with the CAD/IRMPD and CAD/ECD experiments. Charge-reduced species corresponding to the capture of one and two electrons, accompanied by several side-chain losses, can be observed in the mass spectrum in Figure 4.7.f, together with $c$ ions: these are typical ions originated with ion activation by electron-capture dissociation. Many unassigned fragments are also present. They have been investigated using the structure proposed in the scheme of Figure B.1 in the Supporting Information. Direct comparison of the autocorrelation fragment ion scans of Figure 4.7.c and Figure 4.7.f shows the different fragmentation patterns exhibited by the CAD-fragment ${b_{54}}^{5+}$ when subjected to IRMPD or ECD respectively. They demonstrated that CAD/IRMPD and CAD/ECD MS/2D MS experiments are indeed complementary. Horizontal fragment ion scans were extracted for all the successfully assigned precursors, and their fragmentation patterns assigned and used to calculate the cleavage coverage of CaM with the CAD/ECD MS/2D MS experiment. The value of this cleavage coverage is ~32%, calculated considering also the cleavages due to internal fragmentation, and showed in the cleavage coverage map of Figure 4.8.b.

Combination of the data obtained with the CAD/IRMPD and CAD/ECD MS/2D MS experiments of CaM at the detailed condition brought to a cumulative cleavage coverage of ~42%.

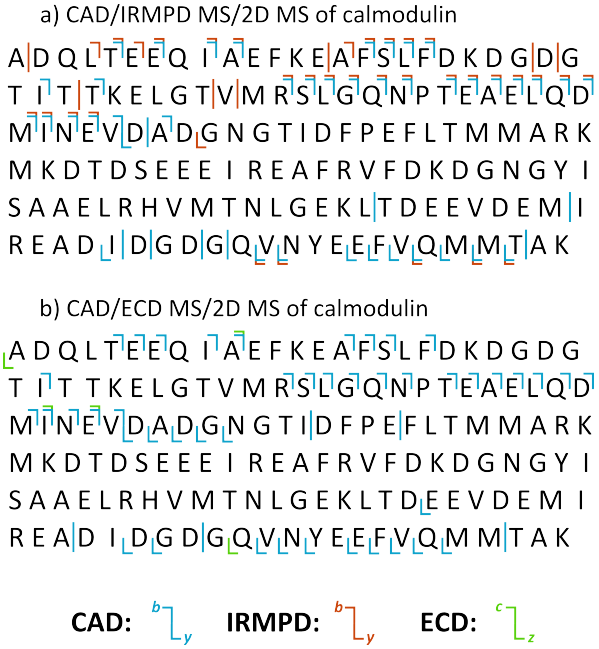



Figure 4.8.: Cleavage coverage maps for the CAD/IRMPD and CAD/ECD analysis of CaM in denaturing conditions. Vertical lines indicate cleavages generated by the formation of internal fragments. The total cleavage coverage for the MS/2D MS analysis of CaM is ~42%.

## 4.4. Discussion

Two-dimensional FT-ICR MS made huge strides in the last decade in terms of processing times and data visualisation: from noisy spectra obtained after several hours of computational processing, to clear ones obtained in a matter of hours or less, thanks to the efficiency of the applied de-noising algorithms and the use of high performance computer clusters. These improvements allowed 2D MS with samples of increased complexity, until the very first applications of the technique to top-down protein analysis, realised on relatively pure samples of Calmodulin using IRMPD as activation method. Top-down 2D MS studies of CaM showed that the technique was able to obtain the fragmentation patterns of all the different charge states of CaM in a single experiment, but the cleavage coverage for the top-down IRMPD 2D MS experiment could be improved. In the work presented in this thesis, a further fragmentation step with CAD, prior to the in-cell fragmentation, is added in order to obtain more structural information from the proteins in analysis.

### 4.4.1. All the $MS^3$-equivalent information in a single experiment

The analyses performed in this thesis showed that with MS/2D MS it is possible to obtain a large amount of information equivalent to an MS/MS experiment on the

entire protein and multiple MS/MS/MS experiments targeted at different fragments in a single run. The MS/MS-equivalent information is included in the autocorrelation line, whose extraction shows a tandem mass spectrum displaying all the fragments obtained through the first fragmentation step (in this case, CAD). Inserts b in Figures 4.3, 4.4, and 4.7 show in fact that extracting the autocorrelation lines in a two-dimensional spectrum from an MS/2D MS experiment a mass spectrum containing typical fragment ion types deriving from CAD is obtained. As mentioned, top-down MS/MS spectra are intrinsically complex (Figure 4.4.b). Extraction and correct interpretation of the autocorrelation line constitutes an important initial step in the top-down MS/2D MS analysis, providing an initial idea of the products from the first fragmentation step. Since a visualisation of the MS/MS spectrum can be obtained in real time before 2D MS acquisition, this stage can be tuned at its best in order to obtain the maximum fragmentation efficiency.

In one-dimensional FT-ICR MS two MS/MS experiments performed with exactly the same parameters might still show slight differences in ions assignments and abundances. In 2D MS these differences are accentuated, because ion relative peak intensity is not tied only to their abundance in the mixture, but also to the fragmentation they undergo in the ICR-cell. That constitutes an explanation of the differences found in the autocorrelation lines of Figures 4.3.b and 4.4.b, although their spectra both represent the first fragmentation of Ubiquitin with CAD, acquired with the same instrument parameters. The cleavage coverage maps of Figure 4.4.d-e show in blue the fragments assigned interpreting the autocorrelation lines of the two MS/2D MS experiments performed on Ubi. In both the spectra generated with the extraction of the autocorrelation lines for the CAD/IRMPD and CAD/ECD MS/2D MS experiments, a percentage of unknown ions is reported. These ions might derive from secondary fragmentation, or from ion "sequence scrambling", a minor ion formation mechanism path that occurs in CAD and IRMPD activations [89]. A substantial difference in the assigned ions can be observed. For both the experiments, the cleavage covered with the analysis of the autocorrelation line was about 45%, unexpectedly more than the initial one-dimensional CAD MS/MS experiment performed as control, whose cleavage coverage ranged around 32%.

An important improvement in the cleavage coverage of the protein is obtained analysing the horizontal scans. For each CAD-fragment found in the autocorrelation line (*i.e.* each precursor in the 2D MS analysis) a fragment ion scan can be extracted, like those shown in inserts c of Figures 4.3 and 4.4 for the ion ${y_{58}}^{9+}$ of Ubi. Fragment ion scans generate mass spectra displaying the fragmentation patterns of the selected precursor. In the case of an MS/2D MS experiment, since the precursors are already fragments derived from an initial CAD activation, spectra derived from their further fragmentation constitute information equivalent to that obtained by multiple MS/MS/MS experiments targeted at those ions. In Figure 4.3.c the horizontal scan

of the ion $y_{58}{}^{9+}$ extracted from the CAD/IRMPD MS/2D MS spectrum shows a wide range of $y$ ions and ions assigned as internal ions. The figure shows also the cleavage coverage map for the spectrum. Further fragmentation of this ion with IRMPD is expected to break more peptide bonds, generating other $b$ and $y$ ions, depending on the terminus that retains the charge, and eventually internal fragments. However, generating a $b$ ion (*i.e.* breaking a peptide bond with charge-retention on the former N-terminus) from a $y$ ion produces what structurally is defined as "internal" fragment, explaining the abundance of internal ions starting with the sequence PSD- present in the spectrum. The fragments found in this spectrum are typical of techniques such as IRMPD, indicating that the ion $y_{58}{}^{9+}$ underwent further fragmentation with such activation method inside the ICR-cell. In the fragment ion scan extracted from the CAD/ECD MS/2D MS spectrum of Ubi, showed in Figure 4.4.c, the generated fragments are much different: $z^{\bullet}/z^{+}$ fragments could be recognised, as well as fragments assigned as $c$ ions, whose theoretical $m/z$ was calculated breaking different $C_{\alpha}$-N bonds along the amino acid chain of the CAD-generated ion $y_{58}{}^{9+}$ of ubiquitin. ECD generated a range of cleavages in the terminal parts of the ion compared to what reported in the cleavage coverage map of Figure 4.4.c for the fragment ion scan of $y_{58}{}^{9+}$ from the CAD/IRMPD experiment.

As comparison, CAD/IRMPD and CAD/ECD $MS^3$ spectra were acquired for the ion $y_{58}{}^{9+}$ of Ubiquitin, reported in Figure 4.7. It is interesting to notice that the cleavages reported for the 1D MS/MS/MS experiments and for the MS/2D MS experiments are different. The extraction of the fragment ion scan from the CAD/IRMPD MS/2D MS experiment presented in Figure 4.3.c reported 19/57 cleavages, while only 4/57 could be observed in the CAD/IRMPD $MS^3$ experiment. The number of cleavages becomes more comparable for the CAD/ECD $MS^3$ and MS/2D MS experiments, reported to be respectively 24/57 and 19/57. While in the first case the MS/2D MS experiment seems to retain more structural information, when ECD is used as second fragmentation technique the number of electron-based cleavages is similar, but the cleavages differ in the highest number of $c$ ions reported for the 1D experiment and for the different cleavage locations, that in the 1D experiment seem to go deep in the ion amino acid chain. Both $z^{\bullet}$ and $z^{+}$ ions are reported for the 1D $MS^3$ and the MS/2D MS experiments.

The difference in the obtained fragments, with consequent difference in cleavages, between the CAD/IRMPD and the CAD/ECD experiments resulted in complementary information, that was summarised in 60/75 cleaved bonds, for a ~83% cleavage coverage. The value is comparable with intensive top-down studies on Ubiquitin that reported cleavages up to 67/75 cleavages using ECD as fragmentation technique [93].

CAD/IRMPD and CAD/ECD MS/2D MS performed on CaM reported a cleavage coverage of respectively ~41 and ~32%, for a total cleavage coverage of ~42%, represented in the cleavage coverage maps of Figure 4.8. Previous top-down 2D MS ex-

periments on CaM in denaturing condition reported a 23% cleavage coverage when IRMPD was used as activation technique [165]. Performing the MS/2D MS experiment on the same protein generated a rise in cleavage coverage of about 19%, representing a further step on the top-down analysis via two-dimensional MS.

### 4.4.2. MS/2D MS for internal ion assignment

Fragmentation techniques such as CAD generate a relevant number of internal ions, from cleavage of both the N- and the C-termini. Such fragments are often difficult to interpret, and their assignment is excluded unless they can be identified with high accuracy. Due to the fact that the breach occurs at the level of the peptide bond for both edges of the amino acid chain, the number of internal ions that might correspond to a single $m/z$ value is often high. Hence the necessity to perform tandem mass spectrometry on internal ions in order to better elucidate their sequence, and correctly assign them. MS/2D MS allows characterisation of all the internal ions that undergo further fragmentation inside the ICR-cell. It is reported the example of the ion of $m/z$ 761.4053, found in the autocorrelation line of the CAD/IRMPD MS/2D MS experiment of Ubi (Figure 4.9). After calibration of the autocorrelation line, the ion's $m/z$ corresponded to two theoretical sequences, PSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLED and EVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQ. Since the atomic composition of these sequences has exactly the same value in $m/z$ at the sixth decimal digit, it would be impossible to discriminate between the two sequences without any sort of MS/MS analysis. Extracting the fragment ion scan for the internal ion of interest from the CAD/IRMPD MS/2D MS spectrum, the spectrum in Figure 4.9 is generated. As expected, all the fragments generated by an internal fragment could be assigned as internals, often with more of a possible sequence per single $m/z$. However, in this case, the presence of the fragment $FAGKQLED^{+}$ ($m/z$ 889.4414) allowed the discrimination between the sequences, confirming the presence of the ion $PSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLED^{5+}$. Although specific fragment ions for the sequence EVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQ were not found, the presence of its 5+ ion cannot be completely excluded, but the assignment of $PSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLED^{5+}$ was important for the cleavage coverage of the entire protein. A table with the complete assignments of the mass spectrum in Figure 4.9 can be found in Supporting Information (Table B.12).

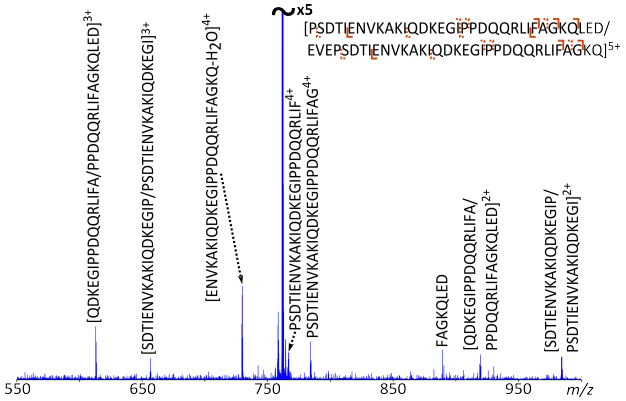



Figure 4.9.: CAD/IRMPD MS/2D MS fragment ion scan for the internal fragment ion of $m/z$ 761.4053, whose sequence could be PSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLED or EVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQ. The presence of the fragment ion FAGKQLED+ ($m/z$ 889.4414) determined the presence of the ion of the former sequence. The cleavage coverage map shows the found (full lines) and hypothetical assignments (dotted lines).

### 4.4.3. Further information available with 2D MS studies

Amongst the key features of two-dimensional mass spectrometry, in MS/2D MS vertical ion scans assume a particularly useful meaning, and support fragment ion assignment. In fact, extracting a vertical ion scan, a mass spectrum of all the precursors that generate a fragment of interest can be obtained. This feature has been shown in top-down 2D MS studies to find all the charge states of a molecule that generated a specific ion [165]. However, in MS/2D MS the precursors are already fragmented ions. Figure 4.3.d shows the mass spectrum generated extracting the precursor ion scan of the ion ${y_{24}}^{4+}$ of Ubiquitin: only $y$ ions are present in the spectrum. A $y$ ion, an ion formed by cleavage of the peptide bond retaining the charge in the N-terminus of the protein, can be generated in-cell via techniques such as IRMPD, SORI-CAD, or UVPD only from other ions containing the C-terminus. In the same way, $b$ ions can be generated in-cell only by other ions retaining the charge in the N-terminus. An example is shown in figure 4.7.d for the CAD/IRMPD MS/2D MS experiment on CaM. Thus, if a prominent, low $m/z$ $b$ ion can be assigned, all precursors that produce that ion (all the vertical precursors in same line) must contain the N-terminus, and similarly a small $y$ ion's precursor scan can identify many C-terminal fragments.

Precursor ion scans were extracted and used for ion assignments for the CAD/IRMPD MS/2D MS experiments of both Ubiquitin and Calmodulin. No significant precursor ion scans were found for the CAD/ECD MS/2D MS experiments. This difference might be due to the different nature of the dissociation techniques used to achieve the second fragmentation: IRMPD is a known ergodic process, which results in selective

cleavages via low energy requirements, while ECD is more random across the amino acid chain [94, 222].

Similarly, neutral-loss lines have the potential to be used for further specific ion assignment if the ions inside the ICR-cell exhibit a specific loss. However, for the data presented here, no convincing neutral-loss lines were found. Unlike for what happens with precursor ion scans, neutral-loss lines are related to the charge of the precursors: this means that precursors with different charge states will be shown in different neutral-loss lines, even if they generate the same loss. Such charge-dependence on the information retained by neutral-loss lines suggests a deconvolution algorithm for two-dimensional spectra, which is currently in progress.

### 4.4.4. 2D MS vs MS/2D MS

Although 2D FT-ICR MS represents the core of the MS/2D MS technique, the parameters for the acquisition of a top-down two-dimensional mass spectrum are different when acquiring it with 2D MS or MS/2D MS. As explained in Chapters 1 and 3, due to the output of 2D FT-ICR MS, which is a two-dimensional map with $m/z$ along both the axis, it is possible to define a vertical resolution and a horizontal resolution [165]. The required resolution could be calculated and obtained based on the complexity of the spectra along its direction. *E.g.*, in a standard top-down experiment the precursors are represented by the different charge states of the protein, which are usually well separated in $m/z$. The fragment ions will instead be disposed in a much more complicated distribution typical of top-down MS/MS experiments. In this conditions, a higher resolution is required for the fragments compared to the precursors. In a top-down MS/2D MS experiments, the precursors are ions derived from a previous (in this case CAD) fragmentation, while the fragments derive from a further fragmentation of these CAD-ions, as shown in inserts c of Figures 4.3, 4.4, and 4.7. Due to the complex nature of top-down MS/MS spectra (showed in the autocorrelation lines of Figures 4.3.b, 4.4.b, and 4.7.b), in MS/2D MS a higher vertical resolution is required compared to the horizontal resolution. Amongst other factors, since the vertical resolution depends on the number of iterations in $t_1$ (scans), MS/2D MS experiments are generally longer than top-down 2D MS experiments.

### 4.4.5. MS/MS/2D MS of Ubiquitin

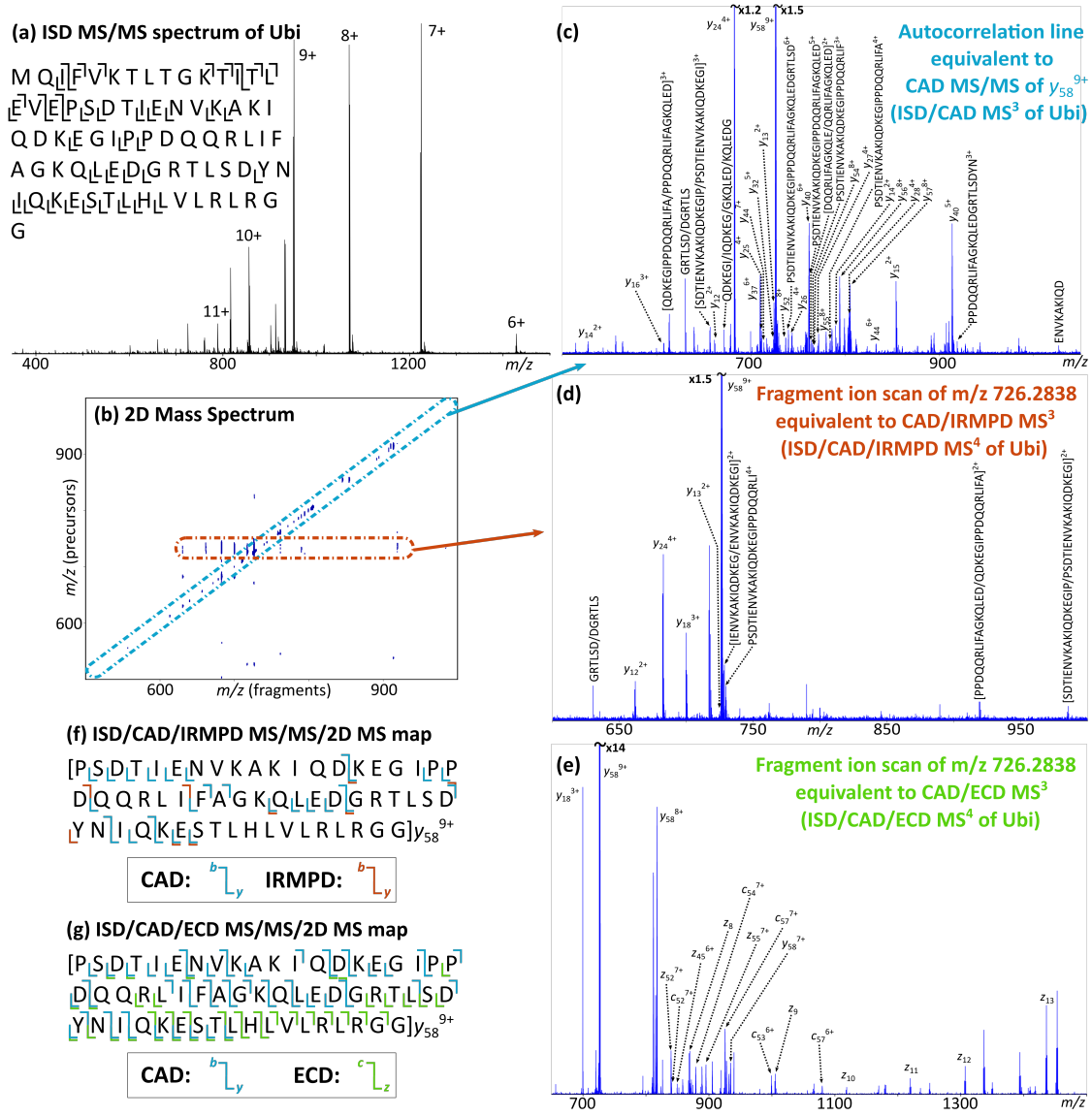



Figure 4.10.: MS/MS/2D MS analysis of Ubi in denaturing conditions. (**a**) ISD MS/MS spectrum of Ubiquitin and respective cleavage coverage map. (**b**) 2D mass spectrum from the ISD/CAD/IRMPD MS/MS/2D MS analysis of Ubi: the autocorrelation line (**c**) and the horizontal fragment ion scan of fragment ion $y_{58}^{9+}$ (**d**) are extracted. (**d**) Fragment ion scan of ion $y_{58}^{9+}$ from an ISD/CAD/ECD MS/MS/2D MS mass spectrum for direct comparison with spectrum d. (**f**,**g**) Cleavage coverage maps for the MS/MS/2D MS analysis of Ubi respectively with IRMPD and ECD as tertiary (in-cell) fragmentation.

As an additional proof of the capability of two-dimensional mass spectrometry for deep sequencing in top-down protein analysis, a further fragmentation step has been added before introducing the fragments into the ICR-cell for 2D MS. For relatively

pure samples, is possible to use nozzle-skimmer dissociations in the source (an in-source dissociation, ISD) to fragment all the charge states of the protein in analysis before isolation in the quadrupole. In the quadrupole, a selected fragment ion is isolated and sent to the hexapole-collision cell for further fragmentation: the generated CAD-fragments are then sent into the ICR-cell for 2D FT-ICR MS analysis, constituting a MS/MS/2D MS experiment. Such experiment is equivalent in information to an $MS^4$ experiment. Nozzle-skimmer dissociations generate the same type of fragments as CAD. The spectrum in Figure 4.10.a shows the initial ISD MS/MS of Ubi with the correspondent cleavage coverage map. Since all the charge states of pure Ubiquitin are fragmented in-source at the same time, peaks corresponding to their $m/z$ are present and labelled in the spectrum. All the unlabelled peaks are $b$, $y$, or internal ions, whose complete assignment can be found in Table B.20 in the Supporting Information. The ion selected for further fragmentation with CAD was the ion ${y_{58}}^{9+}$. Figure 4.10.b shows the two-dimensional mass spectrum obtained with the ISD/CAD/IRMPD MS/MS/2D MS experiment of Ubi with the extraction of the autocorrelation line and a fragment ion scan. The autocorrelation line (Figure 4.10.c) shows the CAD fragmentation of the ion ${y_{58}}^{9+}$ with all the assignments minus the non-diagnostic neutral losses: 31 fragments between $y$ and internal ions can be observed, whose $m/z$ are reported in Table B.21 in the Supporting Information. In MS/MS/2D MS extracting the autocorrelation line gives information equivalent to an $MS^3$ experiment on ubiquitin.

Figure 4.10.d shows the fragment ion scan of ${y_{58}}^{9+}$, a mass spectrum showing the fragmentation pattern of the selected ion under IRMPD. Although CAD and IRMPD generate the same type of fragments, different cleavages are observed due to the different mechanisms of these dissociation techniques. The spectrum in figure 4.10.d is equivalent to a spectrum obtained by ISD/CAD/IRMPD $MS^4$ experiment on ubiquitin. For direct comparison with the mass spectrum in Figure 4.10.d, a fragment ion scan for the ion ${y_{58}}^{9+}$ from an ISD/CAD/ECD MS/MS/2D MS spectrum of Ubi (Figure B.2 in the Supporting Information) is shown in figure 4.10.e. 17 fragments, between $c$, $z$, and $y$ can be observed, typical of ECD fragmentation patterns, and constituting complementary information to the previous spectrum. Similarly to the previous case, the spectrum in Figure 4.10.e constitutes information equivalent to an ISD/CAD/ECD $MS^4$ experiment on Ubi. Fragment ion scans have been extracted for different secondary CAD-fragments from both the MS/MS/2D MS spectra (with IRMPD and ECD), and their complete assignment is reported respectively in Tables B.22 to B.27 and Tables B.29 to B.31 in the Supporting Information. Cleavage coverage maps have been reported for both the MS/MS/2D MS experiments, respectively in Figure 4.10.f for ISD/CAD/IRMPD, and 4.10.g for ISD/CAD/ECD. ISD/CAD/IRMPD MS/MS/2D MS of Ubi gave a cleavage coverage of ~42% for ion ${y_{58}}^{9+}$, and of ~48% for the overall experiment. More promising

was the ISD/CAD/IRMPD MS/2D MS of Ubi, which led to a cleavage coverage of 93% for the ion ${y_{58}}^{9+}$, and of ~87% for the overall experiment.

It is interesting to notice that combining the results of MS/2D MS with only the results of the ISD/CAD/ECD MS/MS/2D MS analysis, a total of 73/75 cleavages is obtained, for a total cleavage coverage of ~97% for ubiquitin in denaturing conditions.

The parameters for the acquisition of these spectra can be found in the supporting information (Table B.32), together with the assignments on the autocorrelation line of the ISD/CAD/ECD MS/MS/2D MS spectrum of Ubi (Table B.28).

### 4.4.6. MS/2D MS with other fragmentation techniques

Other dissociation methods could be used for MS/2D MS. For example, another suitable configuration of the MS/2D MS experiment involves the use of hexapole-based ETD followed by dissociation with IRMPD or UVPD: it would be useful for top-down analysis of phosphorylated proteins, whose first activation with ETD would cleave the protein backbone saving the post translational modifications for the second fragmentation.

### 4.4.7. Towards an improved 2D MS

2D MS is becoming established for top-down analysis of proteins. The computational approaches stepped from long processing and slow visualisation on an office computer to an improved user experience, thanks to the parallelisation of the 2DFT in HPC clusters, the application of more effective denoising algorithms, and the generation of less heavy files for visualisation. Algorithms for a much faster acquisition of high vertical resolution spectra are being developed. The next step in the development of 2D MS will regard the automation of the technique, above all in the data analysis, in order to better manage the huge amount of information obtained.

## 4.5. Conclusions for Chapter 4

A new two-dimensional tandem mass spectrometry acquisition method has been developed, called MS/2D FT-ICR MS or MS/2D MS. The technique is able to obtain information equivalent to a MS/MS/MS analysis of an entire protein, and it has been applied to relatively pure Ubiquitin and Calmodulin in denaturing conditions. Features embedded in 2D MS such as vertical ion scans and neutral-loss lines constitute a different source of information for the sequencing of proteins, and the technique can be developed to $MS^n$/2D MS to deepen the information gathered from intact proteins. The results obtained represent a further step in the top-down "Deep sequencing" of proteins using 2D FT-ICR MS.

# 5. Polymer analysis in the second dimension: preliminary studies for the characterisation of polymers with 2D MS



Available online at https://pubs.acs.org/doi/10.1021/acs.analchem.7b02086.

Two-Dimensional Fourier transform ion cyclotron resonance mass spectrometry (2D FT-ICR MS or 2D MS) allows direct correlation between precursor and fragment ions without isolation prior to fragmentation. The method has been optimised for the analysis of complex mixtures, and used so far for the analysis of small molecules, peptides obtained by tryptic digestion of proteins, and entire proteins. In this work, a 2D MS method is developed to characterise complex mixtures of polymers using IRMPD and ECD as fragmentation techniques, and TPGS, Polysorbate 80, and PMMA as analytes. The use of 2D MS allowed generation of fragment $m/z$ values for all the compounds in the mixture at once, and allowed tandem mass spectrometry of species very close in $m/z$, that would have been difficult to isolate with a quadrupole for standard MS/MS. Furthermore, the use of unique features of 2D MS such as the extraction of neutral-loss lines allowed the successful assignment of peaks from low abundant species that would have been more difficult with standard MS/MS. For all the samples, the amount of information obtained with 2D MS was comparable with what was obtained with multiple 1D MS/MS experiments targeted on each individual component within each mixture but required a single experiment of about 20 to 40 minutes.

## Significance

High-resolution mass spectrometry represents a powerful platform for the analysis of polymers: its use to determine their structural characteristics or their composition, but also to calculate their mass distribution and absolute molar masses, end-groups, and polymerisation mechanisms constitute solid research ground. On the other end, more information can be obtained with the use of tandem mass spectrometry, but such technique is still underdeveloped for polymers. Such condition is due to the high complexity of some of their mixtures (isolation for subsequent fragmentation is not always feasible), to their structure often difficult to fragment (some polymers have considerably stable bonds between monomer units), and to the fact that a different data analysis strategy has to be applied for different polymer classes. Since 2D MS allows fragmentation of all the ions in a mixture without previous isolation, information from all the compounds present in a polymeric mixture can be obtained at once. By setting the instrumentation for the best possible fragmentation conditions, with 2D MS it is possible to see the preferential fragmentation of each compound (horizontal scans), to find species with a similar fragmentation pattern or preferential cleavages (vertical scans and neutral-loss lines), and to have an overview of multiple species at once directly from the two-dimensional map (characteristic diagonal lines). The intent of 2D MS is to provide a new, fast, simplified, data analysis strategy for polymer analysis with mass spectrometry.

## 5.1. Introduction

The increasing diversity and complexity of polymeric structures requires the use of progressively more accurate techniques for their structural investigation. Most polymeric materials consist in fact of complex mixtures of oligomers, whose physical properties are the result of the combined effects of the individual components and are usually characterized through NMR, FTIR and UV spectroscopy, or chromatography. Although these methods have proven their reliability on the studies of polymeric mixtures such as TPGS [223, 224, 225], they struggle to provide a complete profile of the components in the mixture; a difficulty aggravated by the presence of low abundant species. On the other hand, mass spectrometry (MS) is able to characterise the different components of a polymer mixture based on the mass-to-charge ratio ($m/z$) of their ions, through which the number average molecular weight $M_{\mathrm{n}}$ and the weight average molecular weight $M_{\mathrm{w}}$ are calculated.$M_{\mathrm{n}}$ is the total weight of all the polymer molecules in a sample divided by the total number of polymer molecules in the sample. $M_{\mathrm{w}}$ is calculated based on the fraction of the total weight represented by each type of molecule in the sample. The interest in investigating synthetic polymer structures with ultra-high resolution MS strongly increased in the last decade [226].

The high performance of FT-ICR MS/MS has been used to analyse many categories of compounds so far [227, 111], including some categories of polymers. Use of high resolution mass spectrometry to determine structural characteristics of polymers or their composition, to calculate their mass distribution and absolute molar masses [228], end-groups and polymerisation mechanisms [229] are affirmed research domains. An example of end-group determination is represented by the study by Miladinović *et al.* of polyacrylates, in an experiment that attested the essentiality of high mass accuracy and resolution of FT-ICR MS for the study of synthetic polymer complexes [204]. Kaczorowska and Cooper reported the structural characterisation of polyphosphoesters with the analysis of their fragmentation patterns using IRMPD and ECD MS/MS [230]. Finally, FT-ICR MS has been used to analyse the composition of different common polymeric excipients, such as Gelucire 44/14 and Polysorbate 80 [206], and TPGS [205]. In particular for the last case, a detailed ECD and CAD MS/MS characterisation of TPGS was performed, stepping up from an until-then underdeveloped structural characterisation of synthetic polymers due to difficulties in their fragmentation, and the necessity of different approaches for data analysis of different polymers types [204].

In this work, 2D MS is applied for the first time to different polymeric samples in order to develop a method for the characterisation of complex polymeric systems. The samples in use are D-α-tocopheryl polyethylene glycol 1000 succinate (TPGS), polysorbate 80, and poly(methyl methacrylate) (PMMA) (5 kDa).

PMMA is a homopolymer used in a wide range of commodities, and the base for Plexiglas, Acrylite, Lucite, and Perspex. Other than its major use as shatter-resistant glass, so that many pieces of plasticware in a chemistry laboratory itself are made of plexiglass, PMMA is used in many acrylic paint and coating formulations [231]. Recently, FT-ICR MS/MS has been reported for PMMA using both ECD and CAD to highlight the differences in the fragmentation patterns given by the two fragmentation techniques [204].

TPGS and polysorbate 80 consist of polymeric mixtures commonly used as excipients in drug formulation for their ability to improve the bioavailability of poorly water-soluble active ingredients in pharmaceuticals. Water-solubility of drugs is essential for their assimilation in the human body, and polymeric excipients can improve solubility through the formation of micelles in the intestinal tract of patients subjected to the drug of interest. Their addition to drug formulation is, therefore, an important part of the finite pharmaceutical production, and studies of TPGS and polysorbate 80 have been conducted with many different techniques. TPGS consist of a lipophilic head, the tocopheryl component, commonly known as Vitamin E, and a hydrophilic poly(ethylene glycol) (PEG) tail, connected through a succinate diester linker. Other than its mentioned function as absorption/permeation enhancer through the formation of nanoparticles [232], micelles, or liposomes, TPGS is used

as emulsifier and stabiliser thanks to its amphiphilic properties [232, 225]. Interest in the study of TPGS was furtherly developed for its ability to increase bioavailability of anti-cancer drugs by inhibiting P-glycoprotein [224]. Commercially available samples of TPGS contain a mixture of TPGS, Di-TPGS, free PEG and alpha-tocopherol [233, 234, 235]. Studies of TPGS have been performed with NMR, FTIR spectroscopy, UV spectroscopy, chromatography, and standard one-dimensional FT-ICR MS/MS analysis, providing an overview of the polymeric system composition and structural information. However, even if the presence of Di-TPGS has been found in previous studies, its structural characterisation with FT-ICR MS/MS was not performed.

Polysorbate 80 consists of a mixture of sorbitan polyethoxylates, polysorbate diesters, polysorbate monoesthers, and free PEG [207]. Studies of its composition have been performed with IR spectroscopy, chromatography, and mass spectrometry [236, 237, 238]. Combination of liquid chromatography and mass spectrometry (LC MS) improved the initial characterisation of the polymeric mixture [226], but a deeper investigation was performed with FT-ICR MS where the high resolution was essential to overcome spectral complexities, such as the discrimination between species with 0.1 Da difference in mass [206]. However, to date, limited studies about the MS/MS analysis of polysorbate 80 are reported [239].

## 5.2. Experimental section

**Materials.** TPGS was purchased from ISOCHEM, Vert-Le-Petit, France. Polysorbate 80, Poly(methyl methacrylate) (analytical standard for GPC, 5kDa), sodium sulphate ($Na_2SO_4$) were purchased from Sigma-Aldrich (Dorset, UK). HPLC-grade acetonitrile (ACN) was obtained from Fisher Scientific (Loughborough, UK).

**Sample preparation.** The polymer samples were dissolved in acetonitrile, and passed through 0.22 µm centrifugal filters to remove particulates prior to electrospray ionisation. Stock solutions of polymers were diluted to different concentrations (TPGS 5 µM, polysorbate 80 10 µM, PMMA 5 µM) in ACN with 1mM $Na_2SO_4$.

**Instrumentation.** FT-ICR MS was performed on a 12 T Bruker SolariX FT-ICR Mass Spectrometer (Bruker Daltonik GmbH, Bremen, Germany). Samples were ionised using a custom nano electrospray ionisation (nESI) source, using pulled-glass capillaries.. Ions were accumulated in a hexapole for 0.1-0.5 s, then transferred to the ICR-cell, an Infinity Cell [18], for fragmentation, mass analysis, and detection. IRMPD fragmentation was achieved using a continuous wave, 25 W, CO2 laser (Synrad Inc., Washington, USA) held at 50% of its power output. IR photons were

produced at a wavelength of 10.6 µm and pulsed into the ICR cell for 0.28-0.50 s prior to detection. ECD was performed generating electrons from a heated hollow cathode (1.5 A), with a 10 V pulse to inject the electrons into the ICR-cell for 0.03-0.05 s prior to detection (ECD bias 1.5-2.0 V).

**Two-dimensional mass spectra acquisition.** Two-dimensional mass spectrometry was performed using IRMPD and ECD as fragmentation techniques, modulating the fragmentation of the ions according to the scheme in Figure 1.29.a, optimised according to previous studies [178]. The full list of parameters used for the two-dimensional mass spectrometric analysis of polymer samples is provided in the Supporting Information (Table C.15).

**Data processing and analysis.** One-dimensional control spectra were acquired as control, and externally calibrated with the solariXcontrol software (Bruker Daltonics, Billerica, MA, USA) using Agilent ESI-L Low Concentration Tuning Mix (Agilent Technologies, Stockport, UK), then internally calibrated using the theoretical $m/z$ of known ion peaks with a quadratic calibration function in the Bruker DataAnalysis v4.1 software (Bruker Daltonics GmbH, Bremen, Germany).

The 2D mass spectra were processed using SPIKE [192] and the urQRd algorithm for denoising [183]. For the generated spectra, denoising ranks between 10 and 15 were used. The spectra were externally calibrated in the vertical and horizontal dimensions using the same parameters set for the respective one-dimensional control spectra. Internal calibration was performed in the vertical dimension through the extraction of the autocorrelation line and the use of theoretical mass-to-charges of known components, and in the horizontal dimension using theoretical $m/z$ of known IRMPD or ECD fragments.

## 5.3. Results and Discussion

### 5.3.1. Determination of polymer end-groups with mass spectrometry

The codification of a polymer's structure becomes increasingly more difficult the more complex the polymeric mixture. The number of different compounds of a mixture, already enhanced by the distribution of different numbers of repeating units for the same compound, can furthermore increase due to the different adducts that can be formed in solution. Finally, analytical techniques like MS, that use the $m/z$ of the ions in a mixture for analyte identification will see this number multiplied by the number of charges that any single compound can retain. Fortunately, systems like FT-ICR MS have enough resolution to resolve not only the isotopic distribution

of a single compound, determining its charge, but eventual overlapping distributions of different compounds with one or more isotopes at similar mass-to-charge ratios. Knowing the entity of the repeating unit, one of the highest challenges of the structural characterisation of polymers is the determination of the end-group mass. With MS it is possible to determine the end-group mass of a polymer based on the mass-to-charge ratio of its peaks in the spectrum, method used in previous studies of homopolymers and copolymers with MALDI-TOF [240] and FT-ICR MS [241, 242, 243].

Taking into consideration the case of a linear homopolymer, its mass-to-charge ratio on the mass spectrum will be determined by the number of repeating units, its end-group mass, eventual adducts on the molecule, and finally its charge. When the charge is known, the results can be normalised in order to perform the calculations in terms of the mass ($m$) of the macromolecule. The mass of the polymer will be given by the following equation:

$$m_{polymer} = n \cdot m_{monomer} + \sum m_{endgroup} + m_{adduct} \tag{5.1}$$

Where $m_{polymer}$ is the mass of the polymer as it is found in the mass spectrum, $n$ is the number of repeating units or monomers, $m_{\text{monomer}}$ is the mass of a monomer, $m_{\text{end-group}}$ is the end-group mass and $m_{\text{adduct}}$ is the mass of an adduct of the polymer. A polymer in solution can form different kinds of adducts depending on the mixture and the solvent used to solubilise it. For example, samples of TPGS solubilised in 50/50 water/methanol (v/v) showed peaks corresponding to the addiction of $Na^+$, $H^+$, $NH_4{}^+$, and the combination of those adducts, overcomplicating the initial full MS analysis [205]. In MS polymer analysis, experiments tend to be performed preparing the sample with an excess of a specific adduct ion in order for the polymer to be mainly coordinated with the chosen ion instead of forming a more complicated mixture with different adducts. When sample preparation is performed with the addition of a determined ion in the mixture (for example $Na^+$ from $Na_2SO_4$ as described in the method section), the adduct is usually known. When the adduct is known, its mass can be subtracted from the mass of the peak in question, giving the simplified equation:

$$m_{polymer-adduct} = n \cdot m_{monomer} + \sum m_{endgroup} \tag{5.2}$$

The mass of the monomer can also be deducted from the mass spectrum of the polymer, taking in consideration the polymer distribution: the difference between two peaks of the same polymeric distribution (multiplied by the charge of the polymer) gives the mass of the repeating unit. At this point, knowing the number of repeating units, $n$, it would be possible to calculate the missing mass of the end-group with the equation:

$$\sum m_{endgroup} = m_{polymer-adduct} - n \cdot m_{monomer} \tag{5.3}$$

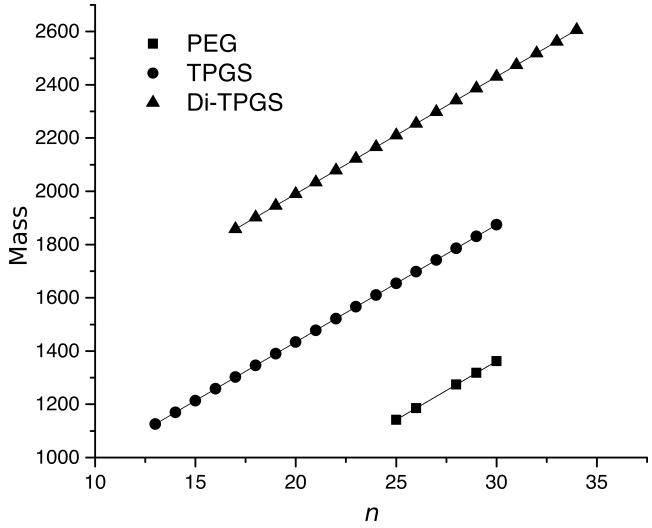
Valid when $m_{monomer} > \sum m_{endgroup}$. However, $n$ cannot be calculated from the peaks of a mass spectrum, and it has to be deducted. Rewriting the equation collecting the mass of the repeating unit, and considering $\frac{m_{polymer-adduct}}{m_{monomer}} = i$, the end-group mass is given by

$$\sum m_{endgroup} = (i - n)m_{monomer} \tag{5.4}$$

Where at this point $n$ is the maximum integer that multiplied by the monomer mass gives the mass of the homopolymer chain, while ($i$-$n$) is a non-integer number that multiplied by $m_{\text{monomer}}$ gives the smallest possible end-group mass, $m_{\text{residue}}$. Once a value for $m_{\text{residue}}$ is obtained through equation 5.4, $\sum m_{endgroup}$ can be calculated and used in Equation 5.1 to determine the consistency of the calculation for a determined polymeric distribution in a full mass spectrum [206].

It is also possible to use the peaks of a known isotopic distribution if $n$ is deducted to graphically calculate the missing addends of Equation 5.1. In fact, plotting the peak masses as they are found in the mass spectrum versus the number of repeating units, $n$, according to equation 5.1 a line is obtained. The slope of the obtained line is given by the mass of the monomers, and the intercept is given by $[m_{endgroup} + m_{adduct}]$.

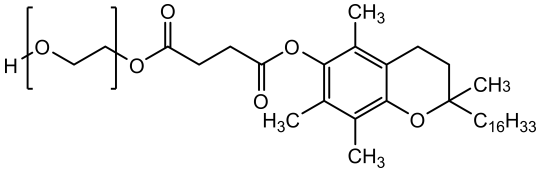
An example is reported for the doubly-charged ions of free PEG, TPGS and Di-TPGS found in the two-dimensional analysis of the TPGS sample using IRMPD as fragmentation mode (see below). The resulting graph plots the convoluted masses of the listed compounds versus the number of their repeating units, $n$ (Figure 5.1). It can be seen from the equation of the linear trends of all the sets of data that the slope is the same for all the lines, averaging 44.0285±0.0026 (mean±SD). The mass of a repetitive unit of PEG is 44.0262 Da, indicating that for all the polymers listed the repetitive unit is ethylene glycol ($C_2H_4O$). The intercepts, determining the sum of the end-group masses and the adducts, indicate for free PEG a $m_{end\text{-}group}$ + $m_{adduct}$ of 40.8415 Da, corresponding to $H_2O$ + $Na^+$; for TPGS a sum of 553.3792 Da, corresponding to $C_{33}H_{54}O_5Na$, the tocopheryl succinate as one end-group, plus a hydrogen as the other end-group, and a sodium adduct; and for Di-TPGS a sum of 1109.7875 Da, corresponding to the two tocopheryl succinate heads plus $C_2H_4$ (see the structure of Di-TPGS, reported in Figure **??**) and a sodium adduct.

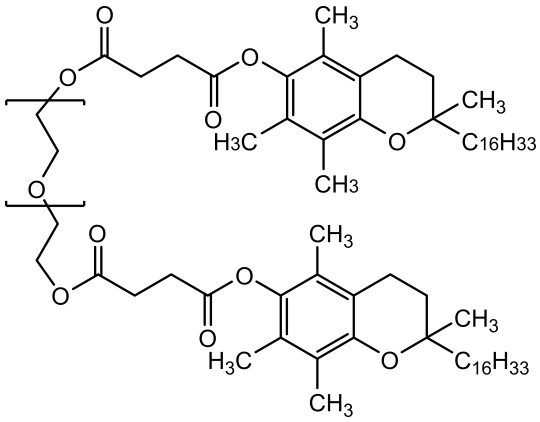
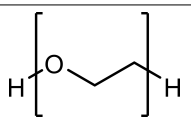
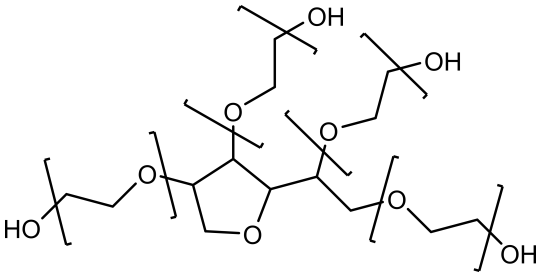
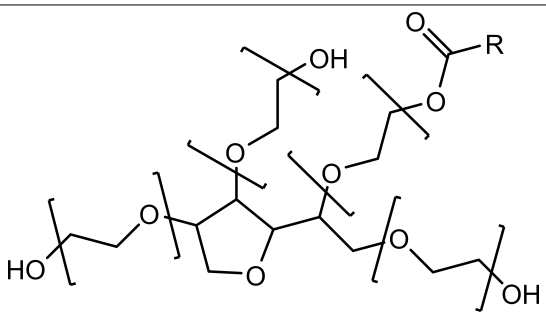
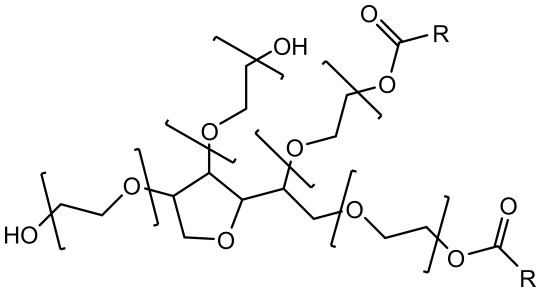
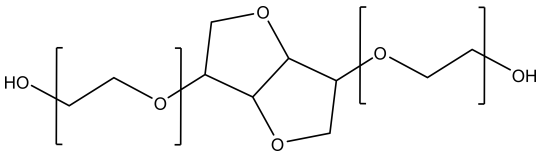


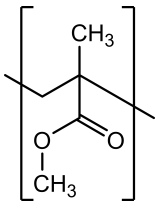

| Equation | y = a + b*x | | |
|---|---|---|---|
| Weight | No Weighting | | |
| Residual Sum of Squares | 6.48834E-4 | 4.22916E-4 | 1.85792E-4 |
| Pearson's r | 1 | 1 | 1 |
| Adj. R-Square | 1 | 1 | 1 |
| | | Value | Standard Error |
| PEG | Intercept | 40.84148 | 0.09809 |
| | Slope | 44.03226 | 0.00355 |
| TPGS | Intercept | 553.37917 | 0.00517 |
| | Slope | 44.0265 | 2.33572E-4 |
| Di-TPGS | Intercept | 1109.78753 | 0.00403 |
| | Slope | 44.02677 | 1.54813E-4 |

Figure 5.1.: Correlation between the mass of the peaks in the autocorrelation line of the IRMPD 2D mass spectrum of TPGS of Figure 5.2 (equivalent to a full mass spectrum of the mixture) and the number of repeating units per mass correspondent to the peaks. The graph, obtained by plotting the data extracted from the 2D mass spectrum for TPGS, Di-TPGS and free PEG with Equation 5.1, shows the consistency of the end-group masses and the monomer masses of the substances in the mixture.

Table 5.1.: Structures of the polymers in each mixture.

| **Mixture** | **Name** | **Structure** |
|---|---|---|
| TPGS | D-α-tocopheryl polyethylene glycol 1000 succinate (TPGS) |  |

| | | |
|---|---|---|
| | Di-TPGS |  |
| | Poly(ethylene glycol) |  |
| Polysorbate 80 | Sorbitan polyethoxylates |  |
| | Polysorbate esters |  |
| | Polysorbate diesters |  |
| | Isosorbide polyethoxylates |  |

| | | |
|---|---|---|
| PMMA | Poly(methyl methacrylate) |  |

### 5.3.2. 2D MS analysis of TPGS

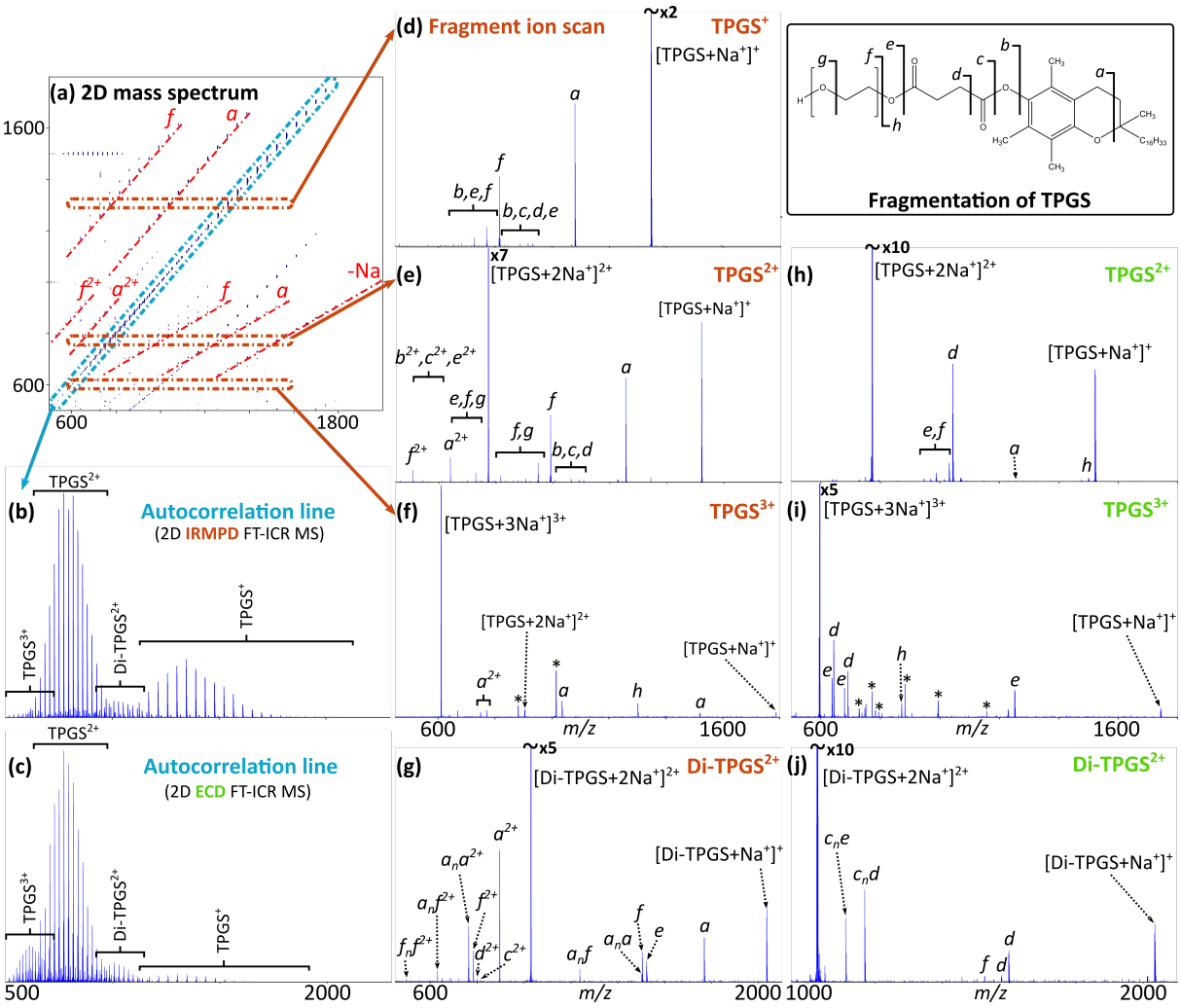



Figure 5.2.: Two-dimensional analysis of the TPGS polymeric mixture in pure acetonitrile and $Na_2SO_4$ with IRMPD and ECD as fragmentation techniques and proposed nomenclature [205]. (a) IRMPD 2D mass spectrum with highlights on the most abundant fragmentations in the mixture. (b) Autocorrelation line extracted from spectrum a. (c) Autocorrelation line extracted from the ECD 2D mass spectrum of TPGS (see Supporting information, Figure C.4). (d-f) Fragment ion scans extracted from the IRMPD 2D mass spectrum of TPGS, respectively for singly-, doubly- and triply-charged TPGS. (g) Fragment ion scan of Di-$TPGS^{2+}$ extracted from spectrum a. (h, i) Fragment ion scans extracted from the ECD 2D mass spectrum of TPGS, respectively for doubly- and triply-charged TPGS. (J) Fragment ion scan of Di-$TPGS^{2+}$ extracted from the ECD 2D mass spectrum of TPGS.

Figure 5.2 shows the results of the two-dimensional mass spectrometry analysis of TPGS performed with IRMPD and ECD as fragmentation techniques. The nomenclature for the fragmentation of TPGS is reported in this figure (top-right) as pro-

posed in previous studies [205]. Figure 5.2.a shows the 2D mass spectrum of TPGS with IRMPD as fragmentation mode. Different characteristic lines can be observed, including a clear autocorrelation line and horizontally distributed peaks showing the fragmentation patterns of all the precursors subjected to infrared multiphoton dissociation. The curved lines departing from the bottom represent harmonics, which do not substantially interfere with the spectral interpretation of extracted horizontal lines due to the unconventional isotopic distribution shown by their peaks. The 2D mass spectrum also shows lines characteristic of this specific polymeric system, indicating the major fragmentation patterns of the main precursors (red -•- lines). The characteristics of the 2D mass spectrum in Figure 5.2.a and the 2D ECD FT-ICR MS spectrum of TPGS are shown in Table 5.2. The horizontal resolving power ($RP_h$), the vertical resolving power ($RP_v$) and the total resolving power ($RP_{2D}$) are calculated according to previous studies [165], and reported at $m/z$ 400. For these experiments the horizontal resolving power was reduced (see the experimental parameters in the Supporting Information, Table C.15) to speed up the overall experimental analysis time, while extending the number of vertical scan lines. The 2D ECD MS spectrum of TPGS is reported in Figure C.4 in the Supporting Information.

| Sample | Dissociation method | $RP_h$ | $RP_v$ | $RP_{2D}$ | $T_{acq}$ [min] |
|---|---|---|---|---|---|
| **TPGS** | IRMPD | 45k | 1.5k | 90M | ~20 |
| | ECD | 28k | 1.6k | 46M | ~20 |
| **Polysorbate 80** | IRMPD | 29k | 1.1k | 32M | ~20 |
| | ECD | 58k | 330 | 19M | ~20 |
| **PMMA** | IRMPD | 55k | 1.1k | 60.5M | ~30 |
| | ECD | 48k | 2.3k | 112M | ~40 |

Table 5.2.: Characteristics of the 2D mass spectra resulting from the analysis of TPGS with IRMPD and ECD 2D MS, and their acquisition times.

The autocorrelation line is extracted in Figure 5.2.b, showing the precursors present in the polymeric mixture. The TPGS sample was dissolved in pure acetonitrile and saturated with $Na_2SO_4$, giving the formation of solely sodium-adduct cations of the analytes in the mixture. Distributions corresponding to different charge states and PEG units of TPGS, Di-TPGS, and free PEG can be observed, with the doubly-charged TPGS being the most abundant species in the mixture and free poly(ethylene) glycol the less abundant species. All the peaks are initially assigned based on their $m/z$ considering $Na^+$ as an adduct, and then confirmed with MS/MS analysis, which in the case of 2D MS corresponds to the extraction of the horizontal lines for the precursors under analysis. The figure reports the distributions of $[TPGS+Na]^+$, $[TPGS+2Na]^{2+}$, $[TPGS+3Na]^{3+}$, and $[Di\text{-}TPGS+2Na^+]^{2+}$, while a list of the complete peak assignments of the autocorrelation line can be found in Table C.1 in the supporting information. Figure 5.2.c reports the extraction of the autocorrelation line from the 2D ECD MS spectrum of TPGS in pure acetonitrile

and 1 mM $Na_2SO_4$. It is interesting to notice that the species found in the mixture are as expected exactly the same, but with a considerable difference in relative abundance. In 2D MS, ion intensity of the autocorrelation line does not depend solely on the ion abundance of the species under analysis. Amongst other factors, fragmentation of the precursors in the ICR-cell determines the precursor intensities on the autocorrelation lines, accentuating the intensity of the ions that fragment more extensively. ECD is known to be a charge-dependent fragmentation, with the more charged species being able to capture more electrons before neutralisation, and consequently to exhibit more fragmentation in their tandem mass spectra. This characteristic of ECD is reflected on the autocorrelation lines of the shown polymeric mass spectra (Figures 5.2.b and 5.2.c), in which the species that retain more charges are relatively more abundant in the autocorrelation line of the 2D ECD mass spectrum than the autocorrelation line extracted from the 2D IRMPD mass spectrum of TPGS. This effect is particularly visible with $TPGS^{3+}$ and $Di\text{-}TPGS^{2+}$, appearing relatively more intense, and of $TPGS^{+}$, relatively less intense in Figure 5.2.c compared to Figure 5.2.b. The assignments on the autocorrelation lines agree with what has been found in previous studies, for which TPGS, Di-TPGS, and PEG were found in a sample of this polymeric mixture using acetonitrile as solvent [205].

The capability of 2D MS to fragment all the species in a mixture without the necessity of single precursor isolation before fragmentation in the ICR-cell is shown with the extraction of fragment ion scans of the main species found in the mixture. Figures 5.2.d. to 5.2.g show in fact the fragment ion scans of respectively $[TPGS{+}Na^{+}]^{+}$, $[TPGS{+}2Na]^{2+}$, $[TPGS{+}3Na]^{3+}$, and $[Di\text{-}TPGS{+}2Na]^{2+}$ extracted from the 2D mass spectrum in Figure 5.2.a. The extracted fragment ion scans show the fragmentation patterns of the selected species subjected to infrared multiphoton dissociation. For a direct comparison between the two dissociation techniques, the fragment ion scans of $[TPGS{+}2Na]^{2+}$, $[TPGS{+}3Na]^{3+}$, and $[Di\text{-}TPGS{+}2Na]^{2+}$ extracted from the 2D ECD MS spectrum of TPGS in pure acetonitrile and 1mM $Na_2SO_4$ are reported respectively in figures 5.2.h to 5.2.j. Figure 5.2.d shows the fragmentation pattern of singly-charged TPGS under IRMPD conditions. The most abundant ion species derive from the fragmentation pathways leading to *a* ions and *f* ions, respectively the break of the hetero-hexane ring at the edge of the tocopheryl head with loss of the long alkane chain, and the complete loss of the tocopheryl fragment leaving only a series of sodiated PEG ions (Figure 5.2, top right). Formation of fragment *a* is intriguing, and it has been hypothesized in previous studies conducted with FT-ICR MS/MS that its formation might be stabilised through the formation of a larger conjugated structure [205], as it happens for Chlorophyll a [115]. *b*, *c*, *d*, and *e* fragments are present in minor abundances. The spectrum shows several internal PEG losses, generating different series of ions with the same termini. Internal PEG losses have been observed in previous studies on samples of long poly(ethylene)

glycol chains using CAD as fragmentation technique [244]. The complete peak assignment of the spectrum in Figure 5.2.d can be found in Table C.2 in the Supporting Information. Previous MS/MS studies conducted with CAD as fragmentation technique did not show any generation of fragments from the isolation and fragmentation of singly-charged TPGS [205]. Such phenomenon might be due to the collisionally deposited vibrational energy detaching the sodium cation preferentially rather than generating other fragments, thus generating only neutral, undetectable fragments. On the other hand, IRMPD at the chosen conditions does not seem to have the same effect when there is only one sodium adduct on the precursor ions, showing the mentioned fragmentation pacts (Fig 5.2.d), with a preference for cleavages in positions *a* and *f*.

The spectrum in Figure 5.2.e shows the fragmentation pattern of a doubly-charged TPGS precursor. As in the case of singly-charged TPGS, two abundant peaks corresponding to *a* and *f* fragment ions can be observed. In the case of the doubly-charged TPGS, however, the most abundant peak seems to correspond to the loss of a sodium cation, generating the $[TPGS+Na]^+$ fragment. Peaks corresponding to doubly-charged *a* and *f* ions can be observed, constituting the third most abundant series of peaks and attesting the preferential cleavages of their termini. Series of *b*, *c*, *d*, *e*, *f*, and *g* ions with different internal PEG losses can be observed at minor intensities, with the *b*, *c*, and *e* ions observed as singly and doubly-charged. The complete peak assignment for the spectrum in Figure 5.2.e can be found in table C.3 in the Supporting information. A fragment ion scan for a $TPGS^{2+}$ peak has been extracted from the 2D ECD MS spectrum of TPGS and reported in Figure 5.2.h: it shows the fragmentation pattern of the selected precursor ion with ECD as fragmentation technique. It can be seen that the two spectra (Figures 5.2.e and 5.2.h) are almost completely different, with the exception of the peak corresponding to the loss of a sodium adduct. Interestingly, one of the major fragmentation pathways in the ECD of TPGS leads to the loss of a sodium cation, and, in this case, no charge-reduced species is observed. The most abundant species in the spectrum is represented by *d* fragment ions, followed by *e* and *f* ions with different internal PEG losses. *d* and *e* fragment ions derive from cleavages close to the carbonyl groups of the tocopheryl head. The presence of an *a* ion is reported at a very low abundance, above all compared with the IRMPD fragment ion scan of the same TPGS species. The presence of *h* ions has been detected, corresponding to the fragmentation of the PEG chain with retention of the charge on the terminus that contains the tocopheryl head. The complete peak assignment for the spectrum in figure 5.2.h can be found in table C.6 of the supporting information.

Spectra in Figures 5.2.f and 5.2.i are generated by the extraction of a fragment ion scan of $[TPGS+3Na]^{3+}$, respectively from the 2D IRMPD MS spectrum and 2D ECD MS spectrum. Figure 5.2.f shows peaks corresponding to the loss of one or

two sodium adducts, and mostly cleavages in position *a*, with formation of different singly- and doubly-charged a ions with different internal PEG losses. For the selected precursor ion, an *h* fragment is observed, but surprisingly no *f* fragment ions were reported. Figure 5.2.i shows a peak corresponding to the loss of two sodium adducts, but no peaks corresponding to a single sodium adduct. As shown in the fragment ion scan of $TPGS^{2+}$, cleavages are preferred in position *d* and *e*, for which respectively two and three peaks are reported, with different internal PEG losses. Even for this spectrum, the presence of *h* fragment ions is detected. It is noticeable that for both the fragment ion scans of the $TPGS^{3+}$ ions an increased number of peaks due to harmonics (indicated with a * symbol) can be recognised. This is due to the fact that these lines have been extracted from precursors in the low part of the 2D mass spectrum (bottom left of Figure 5.2.a), in which noise peaks due to harmonics converge. See supporting information (Tables C.4 and C.7) for complete peak assignment of spectra in figures 5.2.f and 5.2.i.

Previous detailed studies on TPGS using high resolution FT-ICR MS/MS employed CAD, EID, and ECD as fragmentation techniques [205]. Although EID did not show any fragmentation, CAD and ECD were essential for the characterisation of different TPGS mixtures. The fragment ion scans reported in this work for $TPGS^{n+}$ (n = 1, 2, 3) with IRMPD as fragmentation technique contain an amount of information which is comparable to that obtained with the different isolations and fragmentations with CAD reported in the work of Wei *et al.* for the FT-ICR MS/MS analysis of TPGS [205]. Although CAD MS/MS gave more fragmentation of the polymer chain, generating series of fragment ions with the same cleavage and different internal PEG losses, the same cleavages (*a* to *h*) are reported for both the 1D and the 2D FT-ICR MS experiments. However, in 2D MS all the information is retained in a single experiment, with all the species fragmented at once and subsequent saving in time and sample. Furthermore, use of IRMPD to fragment all the species at once allowed the generation of fragments from singly-charged TPGS species, which was not reported for CAD MS/MS in the previous work. The situation is slightly different when comparing the fragment ion scans of $TPGS^{n+}$ extracted from the 2D ECD mass spectrum in Figure C.4 with the ECD MS/MS spectra of the same isolated and fragmented precursors reported in the work of Wei *et al.* Comparing the 1D and 2D ECD MS results of $[TPGS+2Na]^{2+}$, the results are again comparable, showing the same preference for cleavages in *d* and *e* position and the overall presence of *d*, *e*, and *f* ions in both the spectra. The 1D and 2D ECD MS results of $[TPGS+3Na]^{3+}$, however, differ in the number of detected ions, much lower in 2D MS although retaining the information of *d* and *e* as preferential cleavages for ECD of TPGS. The non-detection of some fragment ions reported by Wei *et al.* might be due to the presence of harmonics in the lower part of the 2D mass spectrum, or the low abundance of the fragment ion species in the fragment ion scan. In fact, while in

1D MS transient accumulations are possible in order to increase the signal-to-noise ratio and identify more ion species, spectra generated by extraction of different lines from a 2D spectrum correspond to a single 1D scan.

Finally, a fragment ion scan for doubly-charged Di-TPGS has been extracted from the 2D mass spectrum in Figure 5.2.a, generating the mass spectrum showed in Figure 5.2.g. Figure 5.2.g shows the fragmentation pattern of Di-TPGS$^{2+}$ with IRMPD as fragmentation technique. The structure of Di-TPGS is reported in Table 5.1, and consists of two lypophylic tocopheryl heads connected through a succinate diester linker to a hydrophilic PEG chain. The nomenclature for the cleavages reported in the tocopheryl and succinate components, and in the PEG chain is the same as the one showed in Figure 5.2. As for the case of TPGS$^{2+}$, fragmentation of Di-TPGS$^{2+}$ with IRMPD generates a high intensity peak due to the loss of a sodium adduct. The presence of high intensity peaks corresponding to $a$ and $f$ ions both singly and doubly-charged constitute another similarity with the fragment ion scan of TPGS$^{2+}$ of Figure 5.2.e, showing a strong preference for those cleavage sites. In the spectrum in Figure 5.2.g the presence of an $e$ ion and $c^{2+}$ $d^{2+}$ ions is observed, but none of the mentioned ions exhibits internal PEG losses. Due to its structure containing two tocopherol succinate components and due to the strong preference for cleavages along those parts of the macromolecule instead of the PEG chain, the formation of internal ions with cleavages in both terminal heads is expected. Fragments corresponding to internal ions are indicated with the letter indicating the cleavage position in one of the heads followed by the number of PEG units as subscript, and the letter indicating the cleavage position in the other head of the molecule. *E.g.* the ion $a_{\mathrm{n}}f$ indicates an internal ion of Di-TPGS with a cleavage in position $a$ in one of the heads, $n$ units of PEG, and a cleavage in position $f$ on the other head. Since the molecule is nominally symmetrical (*i.e.* there are no differences between the two heads of the molecule and in the PEG chain, being it a homopolymer) there is no need nor a way to determine which of the two heads is considered the first or the second in the proposed nomenclature. In the spectrum in Figure 5.2.g a series of different internal ions can be observed, including singly and doubly-charged $a_{\mathrm{n}}a$ and $a_{\mathrm{n}}f$, and doubly-charged $f_{\mathrm{n}}f$. Even in this case, no series of ions with internal PEG losses are observed. All the peaks have been assigned with high mass accuracy, and reported in Table C.5 of the supporting information. Finally, a fragment ion scan for a precursor sample of doubly-charged Di-TPGS has been extracted from the 2D ECD spectrum and shown in Figure 5.2.j. As for what is shown in Figure 5.2.h for TPGS$^{2+}$ under the experimental ECD conditions, ECD spectra of Di-TPGS$^{2+}$ generate peaks corresponding to the loss of a sodium adduct and fragment ions with preference for cleavage in the $d$ position. Dissimilarly from the previous case in Figure 5.2.h, although, no $e$ ions are observed. The two $d$ ions reported in the figure differ for the presence of an extra sodium and an electron captured during ECD. No

charge-reduced species is observed, but the presence of an $f$ ion is reported. Peaks corresponding to internal fragments are present, but the favourite cleavages shifted to $c_{\mathrm{n}}d$ and $c_{\mathrm{n}}e$ (see Table C.8 in the supporting information for complete peak assignment). For both the fragment ion scans corresponding to the fragmentation patterns of Di-TPGS$^{2+}$ with IRMPD and ECD, no internal PEG losses have been observed. This is probably easier to understand hypothesising that the internal PEG losses might occur towards the end of the PEG chain, where the end-atom of H might be easier to transfer to the end of the new chain. Such situation is plausible above all with ECD and IRMPD as fragmentation techniques, with mechanisms based on the hydrogen-transfer model [89]. Since in the case of Di-TPGS the end-groups of the polymer are both tocopheryl succinate components, such internal PEG loss might be impeded.

The strong preference for the cleavage in positions $a$ and $f$ for singly- and doubly-charged TPGS, and for the loss of sodium from doubly-charged TPGS and Di-TPGS under IRMPD conditions generate high intensity peaks in correspondence of the $m/z$ of these ions, that can be observed directly on the two-dimensional map. The lines, highlighted in red in the IRMPD 2D mass spectrum, represent the major cleavages for the mixture, providing a mean of characterisation of the polymer in the mixture at a first sight, almost like a fingerprint for polymer mixtures. The lines $a$ and $f$ on the IRMPD 2D mass spectrum of Figure 5.2.a represent signals corresponding to the generation of singly-charged $a$ and $f$ ions from singly-charged (top left) and doubly-charged TPGS (bottom right), and doubly-charged fragments ($a^{2+}$ and $f^{2+}$) from doubly-charged TPGS (bottom left).

### 5.3.3. 2D MS analysis of Polysorbate 80

Figure 5.3 shows the results for the 2D MS analysis of Polysorbate 80 in pure acetonitrile with 1mM $Na_2SO_4$. The resulting IRMPD 2D mass spectrum is shown in Figure 5.3.a (right), with the extraction of the autocorrelation line (5.3.b), two neutral-loss lines (5.3.c and 5.3.d) and a horizontal fragment ion scan (5.3.e).

The IRMPD 2D mass spectrum in Figure 5.3.a exhibits a clear autocorrelation line. Evidence of fragmentation is shown prevalently at $m/z$ lower than the autocorrelation line in the horizontal dimension, indicating at first sight that most of the fragments come from singly-charged ions. Amongst the peaks on the left side of the two-dimensional map, several neutral-loss lines can be observed, and two of them are highlighted (marked with **c** and **d**). Curved lines departing from the bottom part of the spectrum can be observed, constituting as explained harmonic lines. The characteristics of the IRMPD 2D mass spectrum in Figure 5.3.a are shown in Table 5.2.

Analysis of the autocorrelation line is shown in Figure 5.3.b. Several overlapping

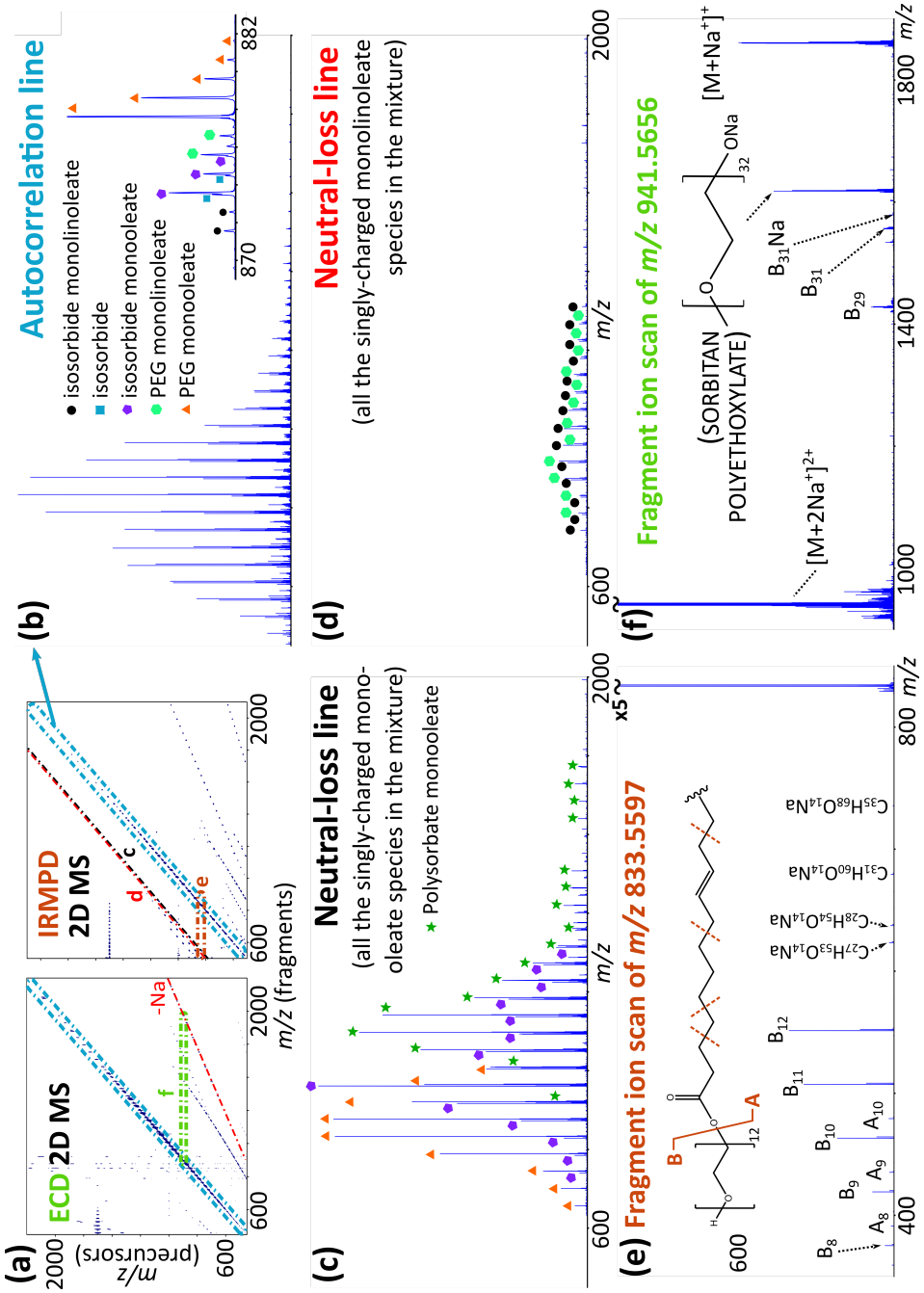



Figure 5.3.: Two-dimensional analysis of Polysorbate 80 in pure acetonitrile and 1 mM $Na_2SO_4$, using IRMPD and ECD as fragmentation technique. (a) IRMPD 2D mass spectrum of Polysorbate 80, with the extraction of the autocorrelation line, two neutral-loss line, and a fragment ion scan. (b) Autocorrelation line extracted from the 2D mass spectrum. (c) Neutral-loss line of all the singly-charged precursors that lost oleic acid during IRMPD in the ICR-cell. (d) Neutral-loss line of all the singly-charged precursors that lost linoleic acid during IRMPD in the ICR-cell. Spectra shown in inserts b, c, and d have the same $m/z$ range, and spectra c and d have the reported at the same intensity scale. (e) Fragment ion scan for $PEG_{13}$ monooleate. (f) Fragment ion scan for the doubly-charged $(PEG)_{32}$ sorbitan polethoxylate monooleate.

distributions can be observed, but the resolving power offered by the FT-ICR technique allows for clear differentiation of very narrow peaks. An example is reported in the insert of the spectrum, showing the array of peaks within about 12 Da. Assignments were made by determination of the end-group masses through the method explained previously [241, 242, 243, 240], and correct $m/z$ matching with the expected species using $Na^+$ as an adduct. Free PEG, sorbitan polyethoxylates, and isosorbide polyetoxylates were found in the autocorrelation line, together with peaks identified through further investigation as corresponding to the same molecules subjected to different kinds of esterification. According to the literature, Polysorbate 80 is the product of a reaction of fatty acids with sorbitol in the presence of an acid catalyst, which generates a mixture of sorbitan polyethoxylates, polysorbate diesters, polysorbate monoesthers, and free PEG (Figure 5.1). The most commonly used fatty acids are oleic ($C_{18}H_{34}O_2$), linoleic ($C_{18}H_{32}O_2$) or stearic ($C_{18}H_{36}O_2$) acid. Based on the $m/z$ of their distributions and further analysis of their fragmentation, polysorbate monooleate, and isosorbate monooleate and monolinooleate were assigned. Unexpectedly, products of the esterification of poly(ethylene glycol) and two of the above mentioned fatty acids were found: PEG monooleate and monolinoleate. The complete assignment of the ions on the autocorrelation line of Figure 5.3.b can be found in Table C.9 in the Supporting information.

Figure 5.3.c shows the neutral-loss line corresponding to the mass of the oleic acid fragment lost as a consequence of infrared multiphoton dissociation in the ICR-cell. The generated mass spectrum shows all the singly-charged precursor ions that lost an oleic acid fragment during the fragmentation period, and hence all the singly-charged precursors esterified with oleic acid. Isosorbide polyethoxilate monooleate, sorbitan polyethoxilate monooleate, and PEG monoleate ions were assigned and are labelled in Figure 5.3. Similarly, Figure 5.3.d shows the neutral loss line corresponding to the loss of a linoleic acid fragment. In this case, the generated mass spectrum shows all the singly-charged precursor ions that lost a linoleic acid fragment during fragmentation with IRMPD, hence all the singly-charged precursors esterified with linoleic acid. Isosorbide polytethoxylate monolinoeleate and PEG monolinoleate were assigned. Diester products were not found in any of the extracted neutral-loss lines. The two neutral-loss lines are reported at the same ion intensity scale to provide a comparison between the relative abundance of the oleated and linoleated species, showing a higher abundance of the monooleated PEG and isosorbide polyethoxylates compared to their linoleated homologous. Extraction of neutral-loss lines like the ones showed in Figures 5.3.c and 5.3.d is a unique feature of 2D FT-ICR MS and it has proven helpful in the ion assignment of the different polymeric components of the analysed mixture. Polysorbate 80 mixture is an example of the products that can derive from polymeric synthesis: when dealing with MS polymer analysis of polymeric mixtures, the number of possible molecules corresponding to a single nominal mass increases

combinatorially, above all considering that, if not the monoisotopic peaks, different isotopes of a compound may interfere with the assignment of another compound because of very similar (even sometimes <0.0001 Da) mass-to-charge ratios. This situation is exacerbated for low abundant species, whose low intensity peaks can be influenced by other peaks' areas, modifying the position of their centroid and inferring with their assignment. It is in the case of such close overlap that, assuming that the ions in question undergo dissociation in the ICR-cell, extraction of neutral-loss lines can help with discerning different peaks. Their capability to be used to check for specific fragment ion loss was particularly helpful in the analysis of this polymeric mixture. As an example, the monooleate ester of sorbitan polyethoxylate with 31 repeating units and two sodium adducts ($m/z$ 919.5524), the monolinoleate ester of poly(ethylene glycol) with 14 units and one sodium adduct ($m/z$ 919.5965), sorbitan polyethoxilate with 37 repeating units and two sodium adducts ($m/z$ 919.5084), and the isotope A+2 of isosorbide polyethoxylate with 17 repeating units and a sodium adduct ($m/z$ 919.4928) are all low abundant species with an average difference of 0.04 Da. In an internally calibrated spectrum with a high signal-to-noise ratio and ultrahigh resolution, using FT-ICR MS to distinguish which peaks correspond to the two found in the spectrum would not be trivial, but as it often happens, above all for preliminary studies of new compounds, mass spectra of unknown mixtures are initially treated with external calibration. Furthermore, the intensity of the ions used for calibration influences the mass measurement accuracy of low intensity ions in the spectra, so in a situation of elevated dynamic ranges, even internal calibration could not guarantee ion assignment with errors under 1 ppm. In the autocorrelation line of the 2D mass spectrum of Figure 5.3.a, two peaks were found at mass-to-charge ratios of 919.5367 and 919.6042. Since $[(\mathrm{PEG})_{17}$ isosorbide polyethoxylate$]^{2+}$ had already been assigned, one of those peaks could have easily been mistaken for being part of its isotopic distribution. The four ions mentioned above all bore errors above 5 ppm when compared to the peaks in the spectrum, so their assignment based only on the $m/z$ of the peaks on the autocorrelation line was challenging. However, extracting the neutral-loss lines corresponding to the loss of oleic acid and linoleic acid from singly and doubly-charged precursors, peaks could be observed at the $m/z$ of $[(\mathrm{PEG})_{31}$ sorbitan polyethoxylate monooleate $+2\mathrm{Na}]^{2+}$ and $[(\mathrm{PEG})_{14}$ monolinoleate $+\mathrm{Na}]^{+}$, giving further information for a successful peak assignment. The neutral-loss line of all the doubly-charged precursors that lost an oleic acid fragment is shown in Figure C.5 in the Supporting Information. The generated mass spectrum shows only the presence of a distribution of sorbitan polyethoxylate monooleate. The intensity of the shown peaks is similar to the intensity of the linoleated precursors shown in Figure 5.3.d.

Fragmentation of the different compounds of Polysorbate 80 polymeric mixture could be examined extracting the horizontal fragment ion scans for a sample of each species.

An example is reported in Figure 5.3.e for $PEG_{12}$ monooleate ($m/z$ 833.5597). The mass spectrum reports the fragmentation pattern of the macromolecule at the experimental conditions of IRMPD. The most abundant fragment ion in the spectrum is represented by the loss of the fragment corresponding to the oleic acid, with the formation of a B ion with all the repeating units of PEG. The nomenclature used for this analysis has been proposed by Cerda *et al.* for the tandem mass spectrometry analysis of PEG, initially performed with ECD [245] and subsequently compared with CAD MS/MS [246]. The dissociation mechanisms of PEG with CAD and ECD are similar, and start with the capture of a hydrogen species ($H^+$ for CAD and $H^\bullet$ for ECD) by an oxygen ether of the homopolymer to trigger a chain cleavage that brings to the formation of A, B or C ions. A ions are generated by the loss of [44($x$-$y$) + (RH)], B ions are generated by the loss of (44$y$+ROH), and C ions represent internal PEG losses, 44($x$-$y$). 44 is the nominal mass of PEG's repetitive units. In the mass spectrum of Figure 5.3.e, five B ions and three A ions with different numbers of repetitive units can be observed. In CAD MS/MS analysis of PEG100, complementary A and B ions could be observed [246], but in this case the $m/z$ of the complementary ions of the A and B ions found in the fragment ion scan would have been too low to be detected. Another interesting case is the formation of ion $B_{12}$ (loss of 44y+ROH), in which $y = 0$ and R = oleic acid. Other than extensive fragmentation of the poly(ethylene glycol) chain, the fragment ion scan showed low intensity peaks corresponding to the fragmentation of the oleic acid chain. One of the most challenging parts of the analysis of different fatty acids regards the position of the double bond. The representation of $PEG_{12}$ monooleate in Figure 5.3.e shows the cleavages assigned for the molecule. The complete peak assignment for the fragment ion scan of $m/z$ 833.5587 can be found in Table C.10 in the supporting information.

Fragment ion scans have been extracted for different compounds assigned to the autocorrelation line. The major peaks found for each ester of PEG, sorbitan polyethoxylate and isosorbide polyethoxylate corresponded to the loss of the fatty acid fragment they were esterified with (oleic acid or linoleic acid).

The ECD 2D mass spectrum, whose characteristics can be found in Table 5.2, is shown in Figure 5.3.a (left). As for the previous case, a clear autocorrelation line can be recognised, but no evident neutral-loss lines or fragment ions in the lower part of the spectrum can be observed, compared to the case of the IRMPD 2D mass spectrum of the same polymeric mixture. Instead, two characteristic diagonal lines can be observed on the low left of the spectrum, one similar to an electron-capture line, and the other representing a preferential fragmentation channel as for the red lines described in the IRMPD 2D mass spectrum of TPGS. Vertical signal strips can be observed for a horizontal mass range between mass-to-charge ratios of 700 and 900. Those lines represent the scintillation noise [180]. The urQRd algorithm included in the processing of the two-dimensional mass spectra of this work (rank 15 for the

spectrum in Figure 5.3.a) helped to minimise it, and, even if it is visible, its presence did not substantially hinder correct mass assignments of the autocorrelation line and fragment ion scans. A curved harmonic line can also be observed at the bottom of the 2D mass spectrum.

The autocorrelation line extracted from the ECD 2D mass spectrum of Polysorbate 80 shows, as expected, *m/z* corresponding to free PEG and isosorbide polyethoxilates, and their oleated and linoleated esters, and sorbitan polyethoxilates and their oleated esters could be found. Comparing the autocorrelation lines from 2D IRMPD and ECD analysis, however, a clear difference in the dynamic range and in the ion intensities can be observed. In the autocorrelation line of the 2D ECD mass spectrum, the doubly-charged sorbitan polyethoxilates and their monooleated esters (the only compounds that have been assigned with more than one charge) are higher in intensity compared to what happens with the 2D IRMPD MS. This phenomenon of intensity magnification on the autocorrelation line of the peaks that fragment the most has been observed and described for the 2D MS of TPGS, and can be observed in this direct comparison of 2D MS analyses of Polysorbate 80 (Figure C.6 in the Supporting Information).

ECD tandem mass spectrometry of Polysorbate 80 has been scrutinised extracting the fragment ion scans of the doubly-charged precursors found. An example is reported in Figure 5.3.f for $[(PEG)_{32}$ sorbitan polyethoxylate monooleate $+2Na]^{2+}$ (*m/z* 941.5656). The generated mass spectrum shows the fragmentation pattern of the selected ion at the experimental ECD conditions. The most abundant peak of the spectrum corresponds to the loss of a sodium cation, with the consequent drop in charge (*m/z* 1860.1419). The ion at *m/z* 1617.8783 is the second most abundant in the spectrum, and corresponds to the fragmentation of the oleic acid between the carbonyl group and the adjacent oxygen, with coordination of an extra atom of sodium. After that, a loss of PEG and water (44y+ROH, where $y = 1$ and R = H) brings to the formation of a $B_{31}$ ion plus a coordination sodium at *m/z* 1555.8401. The ion at *m/z* 1533.8599 counts a loss of about 22 Da, which seems to suggest the loss of the coordination sodium for a proton, with the formation of a $B_{31}$. The explanation for this peak can be found in the structure of sorbitan polyethoxylate esters, reported in Table 5.1. Poly(ethylene glycol) chains are attached to four sites of the sorbitan component: if the fragmentation occurs in the chain where the oleic fragment was previously attached that after dissociation includes an extra oxygen coordinated with sodium the sodiated $B_{n\text{-}1}$ ion is formed. If the dissociation with the loss of PEG+water occurs in one of the other PEG chains, a standard $B_{n\text{-}1}$ ion is formed instead. Finally, in the fragment ion scan of the selected sorbitan polyethoxilate monooleate ion, another ion, $B_{29}$, is observed. The complete peak assignment for the spectrum in Figure 5.3.f can be found in Table C.11 in the supporting information.

Different fragment ion scans have been extracted for sorbitan polyethoxylates monoleate with different number of repetitive units, and all of them reported a peak correspondent to the loss of a sodium adduct, the loss of the oleic fragment with cleavage CO-O and coordination of an extra sodium, and finally the formation of a sodiated or non-sodiated $B_{n\text{-}1}$ ion. The abundance of the peaks corresponding to the loss of sodium and the oleic fragment determine the formation of the above-mentioned diagonal lines on the ECD 2D mass spectrum. The sodium-loss line, indicated in red on the spectrum of Figure 5.3.a (left) and the lettering "-Na" is similar to an electron-capture line, whose equation would be $y = x/(n/z)$ with $n$ the charge of the precursor that lost a charge after electron capture and $z$ is its new charge. In this case, assuming that sodium cations are the only adducts that charge the ions in the mixture, the equation is $y = (x + (n - z)m_{Na^+})/(n/z)$, or $y = (x + m_{Na^+})/2$.

### 5.3.4. 2D MS analysis of PMMA

Finally, a method to analyse polyacrylates with 2D MS has been developed, using poly(methyl methacrylate) as a model. PMMA (5kDa) has been chosen for its reasonable average mass and its homopolymeric nature. 2D MS analysis of PMMA using ECD as the fragmentation technique is shown in Figure 5.4.

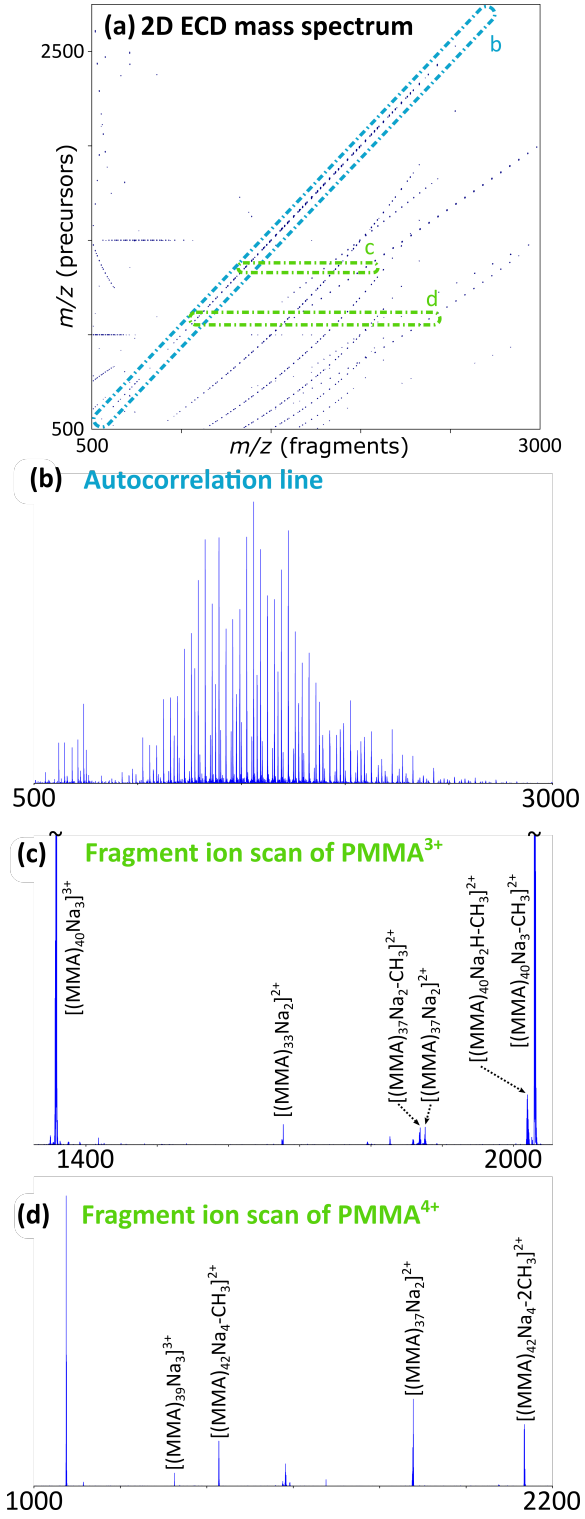



Figure 5.4.: Two-dimensional analysis of PMMA (5 kDa) in pure acetonitrile and 1mM $Na_2SO_4$, using ECD as fragmentation technique. (a) ECD 2D mass spectrum of PMMA. Curved lines are given by harmonics. (b) Autocorrelation line extracted from the 2D mass spectrum. (c,d) Fragment ion scans for the triply- and quadruply-charged ions of PMMA.

Figure 5.4.a shows the resulting ECD 2D mass spectrum, whose characteristics can be found in Table 5.2. As for the previous cases, a clear autocorrelation line can be recognised, together with abundant signals corresponding to intense fragmentations forming characteristic diagonal lines. Several curved lines corresponding to harmonics can be observed departing from the bottom of the spectrum.

The autocorrelation line is extracted and shown in Figure 5.4.b. It exhibits a very complex mixture of PMMA distributions with different charge states, ranging from 1+ to 4+. All the charge states are given by the coordination of $Na^+$ as an adduct in solution. End-groups have been determined for the peaks present in the mixture, reporting only the presence of two hydrogen atoms as termini.

Fragment ion scans have been extracted for $[PMMA + nNa]^{n+}$ ($n$ = 2, 3, 4) in order to study the fragmentation patterns of PMMA with different charge states at the experimental ECD conditions. Figure 5.4.c shows the fragmentation pattern of $[(MMA)_{40}Na_3]^{3+}$ as an example of fragmentation pattern of $PMMA^{3+}$. The most abundant fragment ion peak corresponds to the loss of a radical methyl group from the polymer. This methyl loss is accompanied by a charge reduction of the molecule without any sodium loss. This phenomenon has previously been observed in dissociation studies of cyclic peptides with ECD, and can be explained with the free radical cascade model, a secondary fragmentation path of the used dissociation technique [104]. According to the free radical cascade path, the capture of an electron by the molecule generates a radical cation, which subsequently triggers a series of fragmentations by loss of radicals, such as $CH_3^{\bullet}$. These losses would leave the former molecular ion with all the positively-charged adducts, but with a reduced overall charge. The intermediate charge reduced species has not been observed. The nature of the repeating unit of PMMA (Figure **??**) allows the capture of an electron along the π* orbital of its carbonyl group, as hypothesised in the Utah-Washington mechanism of ECD [247, 103]. Following the electron capture, two possible $-CH_3$ groups can be possibly lost, with the higher probability lying on the methyl group in position alpha with respect to the carbonyl double bond. The second most abundant peak, still several orders of magnitude lower than the peak corresponding to the methyl loss, is a peak in which a sodium atom is replaced by a hydrogen without any further loss in charge. Other peaks in the spectrum correspond to internal MMA losses, reported with two sodium cation adducts, $[(MMA)_{37}Na_2]^{2+}$ ($m/z$ 1874.9622) and $[(MMA)_{33}Na_2]^{2+}$ ($m/z$ 1674.8655). Figure 5.4.d is the mass spectrum generated by the fragment ion scan of a peak of $PMMA^{4+}$, $[(MMA)_{42}Na_4]^{4+}$. The most abundant peak, in this case, corresponds to the break of the poly(methyl acrylate) chain with loss of two sodium cations, $[(MMA)_{37}Na_2]^{2+}$. The other most abundant peaks show the loss of respectively two, for $[(MMA)_{42}Na_4$ $-2CH_3^{\bullet}]^{2+}$, or one, for $[(MMA)_{42}Na_4$ $-CH_3^{\bullet}]^{3+}$, methyl radicals without any loss of sodium cations. A peak correspondent to $[(MMA)_{39}Na_3]^{3+}$ ($m/z$ 1324.3240) is also observed. Complete assignments for the

spectra in figures 5.4.c and 5.4.d are reported in the supporting information (Tables C.12 and C.13). For both the spectra, peaks assigned to experimental noise are indicated with a * symbol. Extraction of fragment ion scans of the doubly-charged precursor reported only the fragment corresponding to the loss of a methyl group.

In 2011, Miladinović *et al.* [204] reported the first study of electro-sprayed polyacrylates analysed with ECD FT-ICR MS/MS. ECD of doubly charged precursor ions of PMMA generated only one fragment ion: a singly-charged ion with the loss of a methyl group. ECD of the equivalent triply charged precursor ions produced the same fragment in both its singly- and doubly-charged forms; it was subsequently determined that the singly-charged products formed via the reaction of an electron with a doubly-charged fragment. The presence of the ions reporting a methyl loss in Figures 5.4.c and 5.4.d can be explained by the free radical cascade dissociation reported by Leymarie *et al.* However, Miladinović *et al.* proposed a similar mechanism. The proposed mechanism for the process involves the capture of an electron by the carbonyl group, followed by cleavage of the $CH_3$ radical and formation of a carboxylate ion. The negative charge of such group contrasts the positive charge retained by the adducts still coordinating the structure, reducing the total charge of the molecule by one, without losing any $Na^+$ [204]. This mechanism would explain the presence of the second most abundant peak in Figure 5.4.c, in which a hydrogen seems to have replaced a sodium atom, previously not observed by Miladinović *et al.* It is hypothesised that a sodium cation is lost and the carboxylate anion is stabilised by acquiring a hydrogen atom during fragmentation. The loss of a sodium cation and neutralisation of the negative charge from the carboxylate anion would leave the total charge of the molecule as it was before this rearrangement, as it is shown in the spectrum of Figure 5.4.c. Another difference observed between the study conducted with standard 1D MS/MS and this two-dimensional study of PMMA is the presence of peaks corresponding to internal MMA losses in the fragmentation patterns of the triply- and quadruply-charged ions obtained from the ECD 2D mass spectrum, absent in the tandem mass spectra shown in the work of Miladinović *et al* [204]. It is hypothesized that isolation of the precursor ions plays an important role in the formation of chain fragments with internal MMA losses because it is the only step not present in 2D MS.

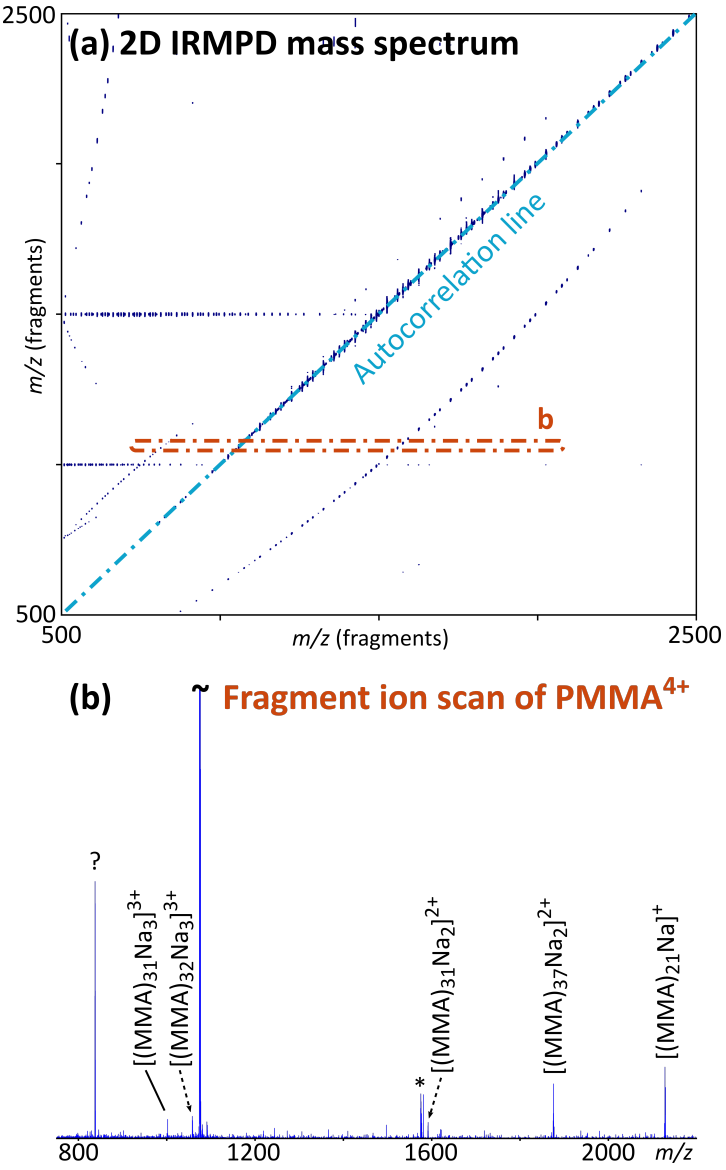



Figure 5.5.: Two-dimensional analysis of PMMA (5 kDa) in pure acetonitrile and 1mM $Na_2SO_4$, using IRMPD as fragmentation technique. (a) IRMPD 2D mass spectrum of PMMA. (b) Fragment ion scans for a quadruply-charged ion of PMMA. Unassigned peaks are indicated with a "?", while experimental noise is indicated with a * symbol.

2D MS has been used to analyse PMMA with IRMPD as fragmentation technique, and the resulting two-dimensional mass spectrum is reported in Figure 5.5.a, together with the extraction of a fragment ion scan 5.5.b. As for the 2D MS analysis, fragment ion scans have been extracted from the 2D mass spectrum, corresponding to a triply-charged precursor, $[(MMA)_{40}Na3]^{3+}$ ($m/z$ 1357.6935) and a quadruply-charged precursor, $[(MMA)_{42}Na_4]^{4+}$ ($m/z$ 1074.0436).

The analysis of the fragment ion scan from $PMMA^{4+}$ shows different peaks corresponding to internal MMA losses, with reduction of the polymer chain length.

PMMA31 was observed both in its triply- and doubly-charged form, while the other ions assigned as PMMA ranged between 1 and 3 positive charges given by their sodium adducts (see Table C.14 for complete peak assignment). In the spectrum, there are also peaks that remain unassigned correspondent to triply-charged ions, at $m/z$ 1080.2121 and 1591.1364, and quadruply-charged ions, at $m/z$ 837.3753 and 837.5025. The triply-charged ion at $m/z$ 1591.1364 is of particular interest because its convoluted mass is higher than the convoluted mass of the precursor ion. Such fact indicates that the precursor or a particular fragment from the precursor underwent molecular reaction during the fragmentation period in the ICR cell. Its $m/z$ is in fact similar to $[(MMA)_{47}Na_3]^{3+}$ ($m/z$ 1591.149173). The presence of quadruply-charged ions indicates the event of a neutral loss or rearrangements amongst fragments that brought to such total number of charges.

The fragment ion scan of $PMMA^{3+}$ exhibited a more curious fragmentation pattern. Although some internal MMA losses were reported, several triply-charged peaks close to the precursor ion were found.

## 5.4. Conclusions for Chapter 5

2D MS has proven to be a powerful and versatile analytical technique for the analysis of proteins and peptides, and, in this work, it is used to develop a method for the analysis of complex polymeric mixtures. Preliminary studies have been performed on two polymeric mixtures commonly used as excipients for pharmaceuticals, TPGS and Polysorbate 80, and on a versatile and ubiquitously-used homopolymer such as PMMA, using IRMPD and ECD 2D MS. 2D MS proved to be suitable for the analysis of polymers, bearing the information of MS/MS experiments in a single two-dimensional run, and offering unique features, such as the extraction of neutral-loss lines, and recognition of preferential fragmentation sites for the polymer compounds through characteristic diagonal lines on the 2D map. The information extracted from the autocorrelation line can be used to determine the polymers' end-groups, and further detailed information is obtained thanks to the simultaneous fragmentation of all the peaks in the mixture, providing information equivalent to MS/MS of all and every precursor in the sample. For all the analysed samples, comparable information is obtained with standard 1D MS/MS with the same fragmentation technique. In the reported case, for TPGS information obtained with 1D CAD MS/MS is comparable with the information obtained by extraction of the fragment ion scans from the IRMPD 2D MS. For Polysorbate 80, these experiments represented the equivalent of the first high resolution tandem mass spectrometry performed on the mixture, due to the fact that with standard 1D MS/MS it would be very difficult to isolate a single species in the quadrupole before fragmentation, because of the complexity of the full mass spectrum.

# 6. Further steps along the two dimensions

The following chapter describes studies whose completion will demonstrate further improvement, development and applications of two-dimensional mass spectrometry.

Being the technique still in development, optimisation of the descrete pulse sequence used to achieve 2D FT-ICR MS is required for different techniques. In this chapter, optimisation of 2D FT-ICR MS for use with electron-capture dissociation is studied, using both a one variable at a time (OVAT) approach and a multivariate study. The technique is subsequently applied to another class of notorious molecules, expanding the use of 2D MS as a powerful analytical tool for glycomics. Finally, development in the hardware for two-dimensional mass spectrometry is explored, presenting the design and realisation of a novel ICR-cell.

## 6.1. Optimisation of two-dimensional mass spectrometry for ECD studies

All the experimental work of this thesis demonstrated that 2D FT-ICR MS is a valuable analytical technique, with several unique features that constitute clear advantages compared to other mass spectrometry techniques. However, being the technique still in development, it presents challenges that need to be overcome in order for it to be enstablished as a routine MS technique. Amongst the main disadvantages of 2D MS, the impossibility to signal-average in the horizontal dimension diminishes the sensitivity of the technique to fragment ions. In 1D FT-ICR MS it is in fact possible to acquire multiple scans, with a signal-to-noise ratio improved by the square root of the number of the scans. This feature is not possible in 2D MS. *I.e.* fragmentation techniques with low fragmentation efficiencies that would produce fragments not readily visible unless the signal is accumulated would not be very effective in 2D MS.

The low fragmentation efficiency of ECD accentuates this problem when the activation is used for TDP. In the MS/MS analysis of entire proteins, in fact, although activation via ECD shows the visible production of charge-reduced species, further fragments will not appear on the spectrum unless the transient is averaged for a larger number of scans (usually between 200 and 500). Furthermore, tuning in 2D MS is different from 1D MS, as explained in Chapter 2, making the fragmentation of the analytes even more challenging. However, the activation technique can still be optimised for an improved fragmentation efficiency, through ion manipulation in the ICR-cell with the pulse programme used to perform 2D MS.

The pulse programme used for 2D FT-ICR MS experiments needs to be optimised in order for the ions in the cell to be modulated, fragmented, and detected using ECD with a hollow cathod [248]. The hollow cathod emits a linear "tube-shaped" electron beam, with no electrons at the center. It has been hypothesized that the highest fragmentation efficency would have been where the area of the beam contains more electrons, which is at a certain radius $r$ from the center of the hollow beam.

To test this hypothesis, a custom pulse program was written, based on the one used to optimise 2D MS with IRMPD [178], in a way that the ions, starting at the center of the cell, were subjected to an excitation pulse prior to fragmentation and detection (see Figure 2.2). The amplitude of the initial pulse was kept at 10.42 Db, while the pulse length was incremented from 0.1 to 4 µs, with increments of 0.1 µs.

The preliminary optimisation was tested on Substance P (SubP) 1 pm/mL in 0.1% formic acid. For each increment of the pulse length, a one-dimensional ECD/FT-ICR MS/MS spectrum was recorded (signal averaged for 20 scans) isolating and fragmenting the doubly charged ion of SubP. The $c$' and $c^{\bullet}$ ions, ion $(a{+}1)_7$, ion $z_9$,

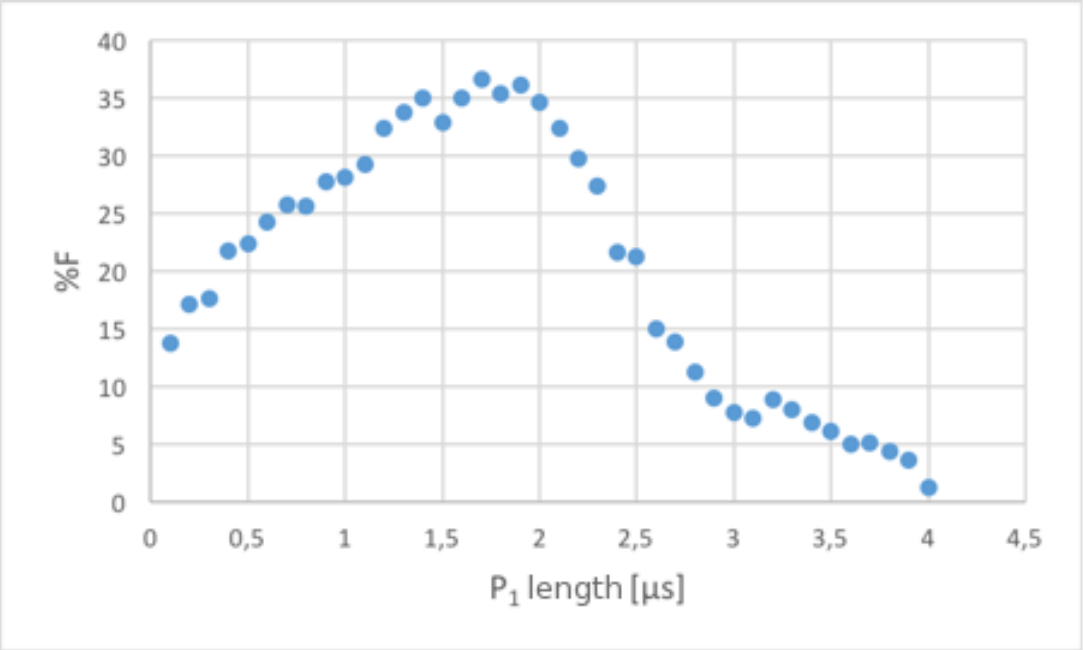



Figure 6.1.: Fragmentation efficiency of SubP with ECD as a function of the pulse-lenght of the initial excitation pulse. It can be seen that there is a maximum at about 1.7 µs, when the ions are supposed to be in the area of the beam more "dense" in electrons.

and the charge reduced species $[MH]^{2+\bullet}$ were used to calculate the fragmentation efficiency, $\%F$ according to the following equation:

$$\%F = \frac{\sum I_f}{\sum I_f + I_p}$$

Where $I_f$ is the intensity of a fragment ion, and $I_p$ is the intensity of the precursor ion.

The results obtained are plotted in Figure 6.1. From the graph, it can be seen that with a small initial pulse the ions are already in the beam area, since a small fragmentation can be observed. With each increment, precursor ions are pushed at high radii until the fragmentation efficiency reaches a point of maximum for an increment of 1.7 µs, then it decreases. The results confirmed the expected ion behaviour, and gave a starting point about the order of magnitude of times and voltages to use in 2D ECD/FT-ICR MS.

Although a general idea on how to tune for 2D ECD was reached, the same values could not be used for the 2D experiments, because the events include a second pulse before fragmentation in order to modulate ion radii. Repeating the same experiment for three points using the standard chirped 2D pulse program (Figure 1.29.a) gave the following results.

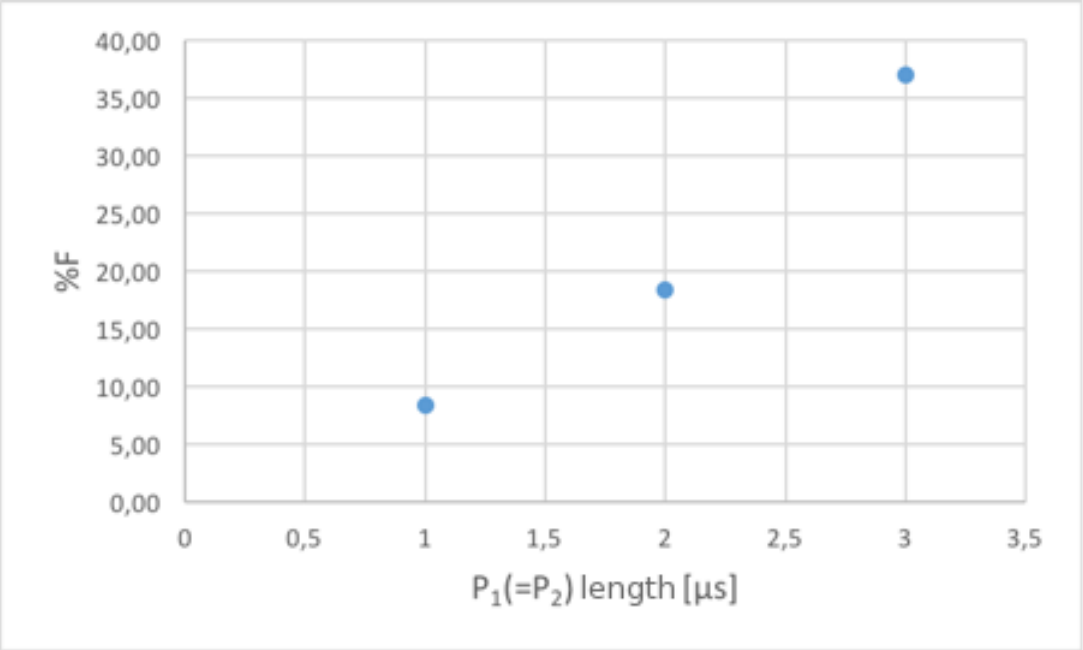



Figure 6.2.: Fragmentation efficiency of SubP as a function of the pulse lenght of $P_1$ and $P_2$ in the 2D ECD/FT-ICR MS experiment. After three seconds, the ions are not at a radius outside of the fragmentation area.

As it can be seen from the graph in Figure 6.2, in case of 2D ECD/FT-ICR MS of Sub P, with a 3 µs pulse, fragmentation efficiency increases instead of decreasing.

To optimise the method, a multivariate analysis has been chosen, varying the pulse length and the amplitude of $P_1$ (and $P_2$), respectively the $X_1$ and $X_2$ variables, and observing how their interaction correlates with the response $Y_1$, the fragmentation efficiency. At the beginning, it was thought to use a $2k$ factorial design (with $k$ number of variables, for a total number of $2^2 = 4$ experiments) as a model, but the use of only 4 experiments was not sufficient to explain the correlations between the two variables. A central composite design, CCD (face centered, with two variables), was used. The CCD allows to estimate the constant, the linear terms, the interactions between variables and the quadratic terms, according to the following model (usually, the interactions among more than two terms are not taken into account):

$$Y = b_0 + b_1X_1 + b_2X_2 + b_{12}X_1X_2 + b_{11}X_1^2 + b_{22}X_2^2$$

Table 6.1 shows the experimental matrix used for the experiment.

Design-Expert® software (version 9, Stat-Ease Inc. Minneapolis, USA) was used to evaluate the quality of the fit to the model, and to calculate the constants and regression coefficient. Results are explained in Figure 6.3 through the use of a response surface. From the surface, it can be seen that the the fragmentation efficiency increases with the pulse length of $P_1$ and $P_2$, and, at a minor rate, with the increase in

| **Pulse length [µs]** | $b_1$ | **Power level (%) [Db]** | $b_2$ | $\%F$ |
|---|---|---|---|---|
| 1 | -1 | (15) 10.42 | -1 | 8.43 |
| 3 | 1 | (15) 10.42 | -1 | 36.98 |
| 1 | -1 | (19) 8.38 | 1 | 15.62 |
| 3 | 1 | (19) 8.38 | 1 | 27.28 |
| 1 | -1 | (17) 9.34 | 0 | 10.84 |
| 3 | 1 | (17) 9.34 | 0 | 26.92 |
| 2 | 0 | (15) 10.42 | -1 | 18.34 |
| 2 | 0 | (19) 8.38 | 1 | 24.79 |
| 2 | 0 | (17) 9.34 | 0 | 18.61 |

Table 6.1.: Experimental matrix used for the optimisation of 2D FT-ICR MS in terms of fragmentation efficiency %F. b1 and b2 represent the factors used to find the coefficients of the model.

excitation power (indicated in Figure 6.3 as percentage power level). It can be seen that the increase in these two variables causes an increase in fragmentation efficiency, but, with a pulse length of 3 µs, increasing the excitation power does not increase the fragmentation efficiency. A local maximum has been found for the fragmentation efficiency, at 3 µs and 10.42 Db (15% of the excitation sweep).

Figure 6.4 shows the Contour plot of the experimental design. The lines connect all the values of pulse length and power level that give the written $\%F$. It confirms that to obtain a fragmentation efficiency higher than 30% for SubP at the ECD fragmentation conditions a pulse length of 3 µs and a power level between 15 and 16% should be used. These results show that after 3 µs at the given excitation powers SubP's ions are not pushed at a radius outside of the electron beam. The design has to be optimised to show an absolute maximum similar to the one showed in the one-dimensional experiment. $\%F = 36.98$ constitutes a local maximum, that will be used as a starting point for the next experimental design.

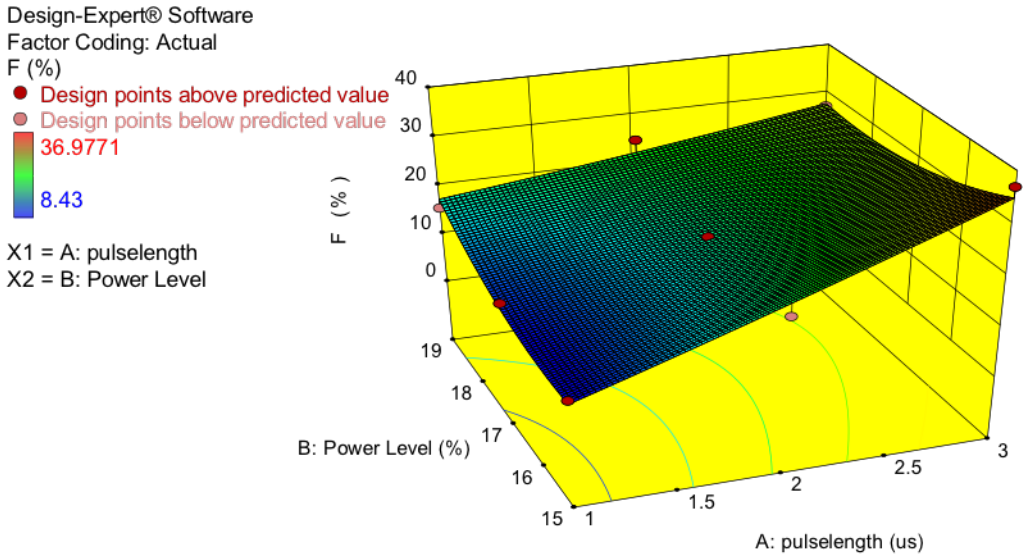



Figure 6.3.: Surface response of the experimental design. It shows the variation of the fragmentation efficiency with the pulse length and the power level of the excitation pulses. A local maximum can be found at 3 µs and 10.42 Db (15% of the excitation sweep).

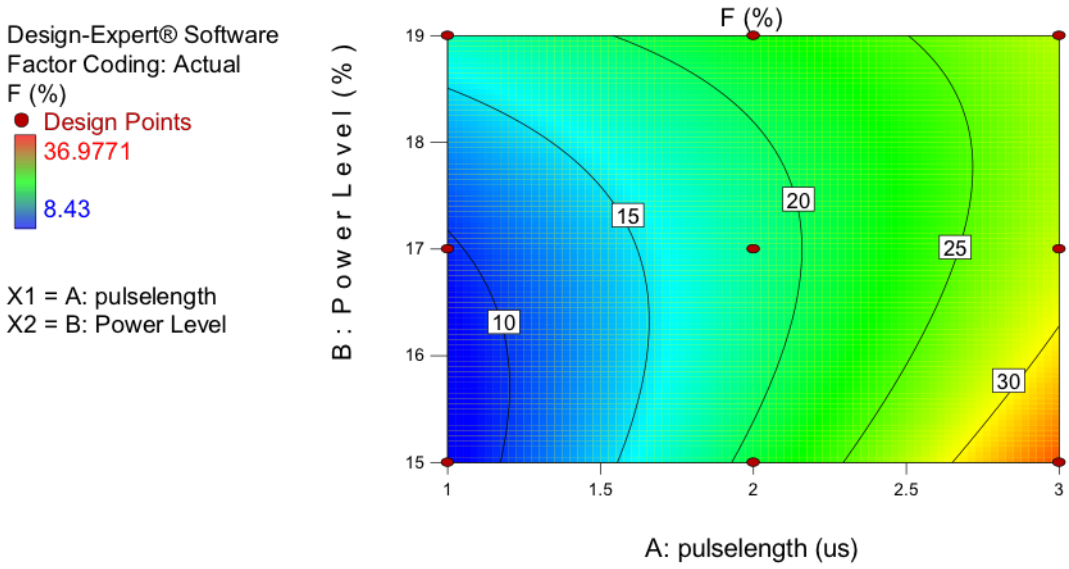



Figure 6.4.: Contour plot of the results. It shows the lines in which the fragmentation efficiency assumes the written value if the corresponding values for pulse length and power level are used.

## 6.2. Two-dimensional mass spectrometry for Glycomics: preliminary studies

Carbohydrates are the most abundant and diverse species present in nature. They occur in nature as heterogeneous mixtures of variable complexity, and are involved in several biological processes, from energy generation to the regulation of biochemical pathways.

Structurally, carbohydrates consist in chains whose the most basic unit is constituted by monosaccharides. Monosaccharides are the simplest form of sugar, the most common of which are the hexoses, with general formula $C_nH_{2n}O_n$, and linear or cyclic forms. The cyclic form of a monosaccharide generates an anomeric center. Typically, the hydroxyl group on the anomeric center is involved in the glycosidic linkage that implicates the connection of a monosaccharide unit to another residue. Glycosidic linkages are defined as α or β based on the direction of the hydroxyl group involved in the bond with respect to the plane of the ring. Modifications of monosaccharides include the loss of one or more oxygen atoms with formation of deoxyacids, oxidation of a hydroxyl group to carboxylic acid to form uronic acid, and amination of a hydroxyl group, replaced by a primary amine or an acylaminogroup. Chains of two to ten monosaccharide units constitute oligosaccharides, while over ten units the resulting macromolecules are called polysaccharides. Both oligo- and polysaccharides are made by the attachment of monosaccharides linked to one another via glycosidic linkages, and can be observed as linear or branched chains. Polysaccharides are typically composed by repeating oligosaccharide motifs. Numerous biological processes derive from specific carbohydrate structures [249, 250, 251].

The importance of carbohydrates in the regulation of different processes in the human body has been accentuated by the human genome project, in which the number of identified genes was much lower than the estimated value, leading to the conclusion that much of the complexity in the human body arises from protein differentiation rather than a high number of various genes. On the protein level, variation can arise from many post-translational modifications (PTM's), the most common of which is glycosylation [252]: it is estimated that nearly half of the proteins fundamental to living beings are glycosilated [253]. Glycosylation is the process through which one or more glycans (mono- or oligosaccharides) are attached to peptides, proteins, or lipids. The attachment of a mixture of glycans to a protein modifies its structural properties, including solubility and conformation. Glycosylation of the lipids forming the external barrier of a cell influence cellular signalling and cell adhesion [253, 254, 255, 256, 257].

The attachment of glycans in specific sites of a protein generates structurally and chemically diverse glycoproteins. Glycoproteins often exist as mixtures of proteins sharing the same aminoacid sequence but different glycan structures and allocations:

these sub-groups are called glycoforms. The most common forms of glycan-protein bonds are the *N*-glycosylation and the *O*-glycosylation. In *N*-glycosylation, the glycan is attached to the nitrogen atom of an asparagine side chain (from which the name), if the protein contains the consensus sequence N-X-S/T, where N is Asparagine (or Asn), S is Serine (or Ser), T is Threonine (or Thr), and X is any aminoacid except Proline (Pro, or P). The *N*-glycans of upper organisms all contain a core structure called trimannosyl chibitose core [258], a pentasaccaride from which other oligosaccharide chains can branch. The trimmannosyl chibitose core is made of three units of mannose (Man) and two units of N-acetyl glucosamine (GlcNAc), $Man_3GlcNAc_2$. In *O*-glycosylation, the *O*-glycans are attached to the hydroxil group of a Ser or Thr, and they do not have a consensus sequence nor a consistent core structure, facts that enhance their diversity and the challenge of their analysis [259, 260].

The study of glycans is called glycomics, and implicates the structural characterisation of the involved macromolecules, their function, and the distribution of glycoforms at different levels (cellular, tissue, organ, and organism). Opposite to what can be applied to other macromolecules (such as overexpression for proteins or polymerase chain reaction, PCR, for nucleic acids), there is no method to amplify the amount of carbohydrates present in a sample, limiting the bioanalysis to what is obtained from natural sources. Furthermore, unlike for what happens to other classes of biomolecules, a small number of monosaccharides can easily form many stereoisomers, quickly increasing the complexity of their analysis. As a matter of fact, even the smallest differences in isomerism or chirality may bring different biological activity. Mass spectrometry (MS) constitutes an important platform for glycomics, due to its high sensitivity and its ability to deal with complex mixtures. Characterisation of glycans or glycoproteins through MS is based on obtaining all the structural information from the mass spectrum, involving tandem mass spectrometry (MS/MS or $MS^n$) of isolated species with different fragmentation techniques, and sometimes chromatographic separations and chemical derivatisation, for major elucidation.

In order to characterise the carbohydrate profile of a glycoprotein, glycans need to be released from the protein moiety. Two general methodologies are available, enzymatic release and chemical release. Enzymatic methods are preferred for the advatage of retaining the protein structure for further analysis. *N*-glycans are released from glycoproteins through the use of specifc enzymes, endoglycosidases, which remove most of the glycan from the protein structure, and exoglycosidases, which remove specific parts of the glycan from the non-reducing end. The most common family of endoglycosidases for *N*-glycan release are the protein *N*-glycosidases (PNGase). PNGase F releases *N*-glycans by breaking the bond between the first GlcNAc of the trimannosyl chibitose core and the Asp, leaving the aminogroup on the sugar intact, which altough quickly hydrolises to be substituted by an hydroxyl group [261, 262],

while PNGase F1-F3 leave the innermost GlcNAc attached to the protein. The latter can be used when more information is needed for specific allocation of the *N*-glycans on the glycoprotein. Chemical release involves alkaline elimination and hydrazinolysis [263, 264, 265], and it is frequently used for *O*-glycan release.

Released glycans have a reducing end that can be derivatised in order to improve stability, ionisation efficiency, and to add a stable isotope label or a chromophore for UV detection (following high-performance liquid chromatography, HPLC) [266, 267, 268]. Another important derivatisation for the characterisation of glycans is chemical methylation (permethylation) of their -OH, -$NH_2$, and -COOH groups. Permethyation allows uniform ionisation for both acid and basic glycans [266], decreases the intermolecular hydrogen bonding, enhancing volatility and thus signal intensity [269], and prevents gas-phase rearrangements [270]. The substitution of hydrogen atoms from electronegative atoms with a methyl group reduces the polarity of the molecules, allowing more reproducible LC runs. Finally, permethylation determines the formation of diagnostic ions in MS/MS, yielding fine structural information [269, 271, 272, 273, 274, 275, 276].

The most widely used techniques for the ionisation of carbohydrates are matrix assisted laser desorption/ionisation (MALDI) [25] and electrospray ionisation (ESI) [31]. Both are defined as *soft* ionisation techniques, for the low energy that they impart to the analytes, generating mostly molecular ions with little fragmentation. In the case of glycans, loss of the most labile groups, especially sialic acid and fucose can occur. MALDI generates a series of singly-charged ions, with the ionisation efficiency of neutral analytes being independent from their molecular size. The ionisation efficiency of ESI decreases with the increase in molecular weight of the analytes [277]. Furthermore, ESI generates multiply charged ions, feature that might complicate spectral analysis of complex mixtures, but increases the fragmentation efficiencies of charge-dependent fragmentation techniques such as the electron-based dissociations (ExD).

Amongst all the mass spectrometric techniques, FT-ICR MS offers the highest performance in terms of resolving power (RP) and mass measurement accuracy (MMA), reaching routinely RP $> 10^6$ and MMA $< 1$ ppm with 12 T magnets and ultra-high resolution. The versatility of FT-ICR MS is accentuated by the different fragmentation techniques that can be associated with such instrumentation, offering a wide range of possibilities for tandem mass spectrometry. Some ExDs, for example, such as ECD and EDD are prerogative of FT-ICR MS, and have been used to investigate linear and branched oligosaccharides. Extensive cross-link cleavages of metal adducts of glycans with ECD were demonstrated by the works of Zhao *et al.* and Adamson and Håkansson [278, 279], and Wolff *et al.* elucidated some fragmentation mechanisms of glycosaminoglycans (GAGs) through the use of EDD [280]. Adamson and Håkansson obtained relatively abundant cross-ring cleavages of negatively

charged N-glycans and milk oligosaccharides [281]. Yu *et al.* used different ExDs, ECD, hot-ECD, and EID (see Chapter 1) to characterise an alkalin metal-adducted permethilated maltoheptaose, allowing for detailed characterisation of the analyte [282]. Low-energy activation techniques such as CAD and IRMPD produce mostly glycosidic cleavages, but cross-ring cleavages can be increased by a metal ion binding to a ring oxygen [283, 284].

With its capability to retain post-translational modifications during fragmentation of proteins, ECD can be used to sequence proteins and peptides without affecting their glycosylation sites. IRMPD is able to "knock off" the linked glycans or part of them from the protein/peptide under analysis, and to produce fragment ions of both the protein and the glycans, depending on the required power. Both the activation techniques are performed in the ICR-cell and are compatible with 2D FT-ICR MS. Given a glycoprotein, it would be possible to identify the glycosylation sites and sequence the structure with two BUP 2D MS experiments, respectively with IRMPD and ECD. The loss of a glycan or a glycan unit could be identified for each peptide in the mixture through the extraction of neutral loss lines, allowing to easily recognise all the glycosilated peptides. The same peptides could be all sequenced in a single 2D ECD MS experiment. Furthermore, if glycans are extracted from the protein digest, they can be analysed all at once with 2D MS.

Following these hypotheses, samples of RNAse B and fetuin were digested and subjected to 2D MS analysis, with the goal of identifying their glycoforms with the smallest number of analyses using the features of 2D FT-ICR MS. RNAse B is a high-mannose *N*-linked glycoprotein often used as a model for similar PTM studies, while fetuin presents several sites of *O*-glycosylation.

### 6.2.1. Experimental section

**Materials.** RNaseB, iodoacetamide, dithiothreitol, and NaOH were purchased from Sigma-Aldrich (Dorset, UK). Fetuin and PNGase were purchased from Promega Corporation (Madison, USA). HPLC grade methanol and formic acid (HAc), were obtained from Fisher Scientific (Loughborough, UK). Water was purified by a Millipore Direct-Q purification system (Millipore, Nottingham, UK). The kit used for the tryptic digestion of RNase B and Fetuin was purchased from Thermo Scientific (Loughborough, UK).

**Sample preparation.** Aqueous stock solutions of RNaseB and Fetuin were prepared. Aliquotes of the proteins were digested using the Thermo Scienfic SMART digest kit according to the manual, performing the digestion for a X h at Y °C and desalting the peptides using μ-SOLA SPE cartridges, washed twice with acidified water, then

eluted with 80% ACN/water. For fetuin, different elution steps were performed, using respectively 10%, 20% 60% and 80% of acetonitrile in pure water. The resulting peptides were subsequently reduced and alkylated using respectively dithiothreitol (DTT) and iodoacetamide (IAA). RNaseB (5.8 µM) was dissolved in a 75:25 water/acetonitrile (v/v) solution with 0.3% (v/v) of formic acid. Fetuin (6.0 µM) was dissolved in a 75:25 water/acetonitrile (v/v) solution with 0.3% (v/v) of formic acid.

**Mass spectrometry.** The mass spectrometry analysis was performed using a solariX FT-ICR mass spectrometer equipped with a custom nanoelectrospray source, and IRMPD and ECD capabilities, as described in previous chapters. One-dimensional control spectra and two-dimensional mass spectra were acquired for the tryptic digests of RNAse B and fetuin, using ECD and IRMPD as fragmentation techniques.

### 6.2.2. Preliminary results and discussion

The samples of RNAse B and fetuin have been prepared, obtaining for each protein a mixture of glycopeptides and a mixture of detached glycans. Two-dimensional mass spectrometry analysis was performed on the glycopeptide mixtures, and the results are under further investigation.

Figure 6.5 shows the 2D mass spectrum resulting from the analysis of RNAse B with IRMPD as a fragmentation technique. By extracting the autocorrelation line, the mixture of glycopeptides is elucidated, showing a wide array of enzymatic cleavages, in a mass range between $m/z$ 500 and 1500. The figure shows a prominent feature of 2D MS for the analysis of glycans: the extraction of a neutral-loss line, in this case showing all the doubly-charged precursor ions that lost a mannose unit during IRMPD. This feature allows to indicate immediately the peptides with the post-translational modification.

The spectrum reported in Figure 6.5 represents the first glycoprotein tryptic digest to be analysed through two-dimensional mass spectrometry. At the moment, further studies are in progress to realise glycoform assignment through 2D MS, using RNAse B and fetuin as models.

Figure 6.5.: Two-dimensional IRMPD mass spectrum of tryptic digested RNAse B. Its acquisition took about 30 minutes. The autocorrelation line is extracted, showing the mixture of BUP peptides, as well as a neutral-loss line of all the doubly-charged peptides that lost a mannose unit during fragmentation in the ICR-cell. The purple circles indicate the glycosolated peptides found through extraction of the neutral-loss line.

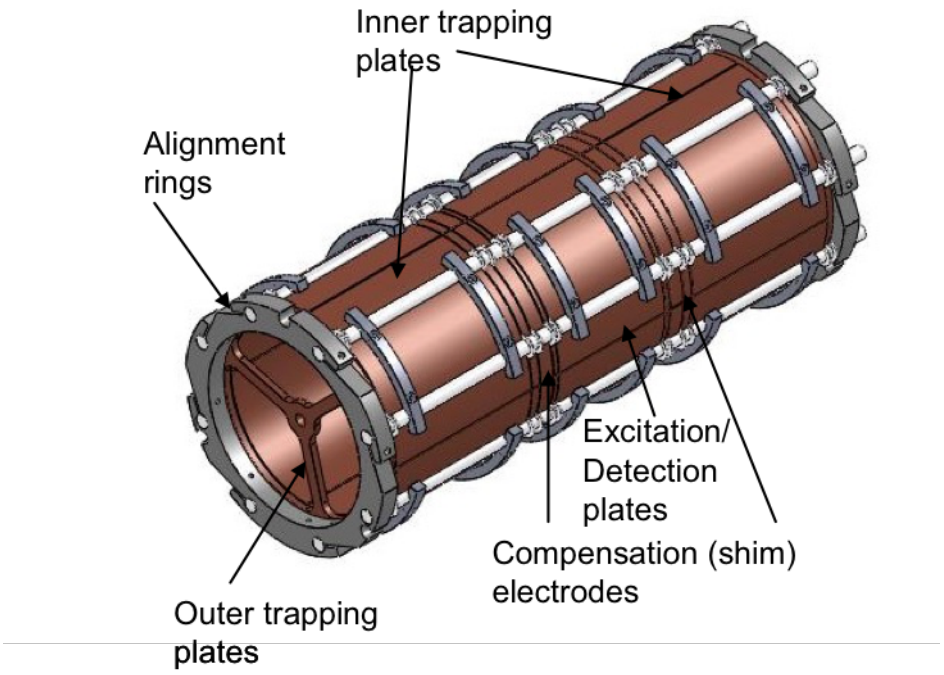



Figure 6.6.: Schematic representation of the newly designed compensated cylindrical cell. The cell is constituted by four quarters ("sides") forming a cylinder. Each side has two external inner trapping rings, a central detection or excitation plate alternating along the cylinder diameter, and a pair of three compensation electrodes between the trapping and the detection/excitation plates. Each plate is separated by ceramic washers.

## 6.3. A novel ICR-cell for two-dimensional mass spectrometry

A novel ICR-cell has been designed and built in order to accommodate the needs of 2D MS, and FT-ICR MS in general, of a better ion population capacity, higher performances in terms of RP and MMA, and a simpler tuning procedure. Figure 6.6 shows the design of the ion trap, an open compensated cylindrical cell. As explained in Chapter 1 (see Sections 1.2.1, "Principles of FT-ICR MS", and 1.4, "ICR cells"), the ICR-cell is the core of the FT-ICR mass spectrometer, and it is enclosed inside a superconductive magnet bore in order to guarantee the ion motions at the base of FT-ICR MS theory. The cylindrical shape allows a perfect enclosing inside the magnet bore, hence a good linearity of the magnetic fields inside the cell.

As for the design by Brustkern *et al.* [64], in the new ICR-cell a pair of three compensation rings ("shim" electrodes) is located between the excitation or detection plates and the trapping plates. A series of structural choices have been made in order to improve the design. The diameter of the cell is larger compared to other cylindrical

Figure 6.7.: Different gold coatings on aluminium samples. The left part of the figure shows the coating of a piece of aluminium with pure gold: the surface was not smooth. The sample on the right, a ring of aluminium coated with a layer of nickel and one of gold, shows a smooth, consistent surface.

cells. Its outer diameter of about 10.4 cm allows the cell to take full advantage of the diameter of the magnet bore. An increase in diameter results in a decrease in the radial component of the force acting on ions because of the electrostatic trapping potential. This larger geometry should reduce the field inhomogeneity experienced by the ions, improving the achievable mass measurement accuracies. A larger diameter also allows the cell to contain more ions before the ion clouds start to be affected by the space-charge effects. In this way, longer transients can be recorded, resulting in higher resolving powers.

### 6.3.1. Materials

The cell plates have been realized using wire electrical discharge machining (wire EDM) by the engineering department of the University of Cambridge. Wire EDM involves a series of rapid electrical discharges between a micrometre-sized wire and the aluminium material, causing micrometre sized pieces of the aluminium to erode away and creating the desired plate geometries. This technique ensures that the tolerances of the pieces are kept below 10 μm so that the plates fit together as a cylinder with a very high degree of accuracy. This will ensure that the electric field created by the trapping potentials is not perturbed by misalignment of the electrodes. The cell plates are entirely made of aluminium, material chosen for its paramagnetism, hence the capability to be inserted in a high magnetic field without interfering with it, and with the ion motions. However, aluminiun tends to oxidise with time (even in UHV), and the layer of aluminium oxides can influence the response of the electrodes to the applied voltages, leading to poor ion motion control and reduced performance. A smooth, consistent surface on the inside of the cell is needed for a correct ion manipulation. In order to avoid oxidation, the surfaces of each electrode have to be coated with a layer of a noble metal. A gold-coating has been tested on an aluminium sample, but the coating created a coarse surface, unapplicable to the purposes of the ICR-cell, as shown in the left part of Figure 6.7. In order to prevent this phenomenon, a layer of nickel is applied between the gold and the aluminium. The result is shown in the right part of the figure.

As the cell has to be installed in a 12 T FT-ICR mass spectrometer, all the materials

| Sample | M [emu/g] | H [kOe] | χ [$cm^3$/g] |
|---|---|---|---|
| **Al** | 0.002 | 10 | 2.51 x $10^{-6}$ |
| | 0.017 | 70 | 2.51 x $10^{-5}$ |
| **Al/Ni/Au** | 0.020 | 10 | 1.88 x $10^{-5}$ |
| | 0.042 | 70 | 3.05 x $10^{-6}$ |

Table 6.2.: Magnetic susceptibilities calculated at 10 and 70 kOe, approximately 1 and 7 T for a sample of aluminium, and one of gold/nickel-coated aluminium.

used for the plates have to be chosen in a way that the magnetic field does not magnetize the cell, inducing a secondary magnetic field which would interfere with ion motion. Magnetic susceptibility, $\chi$, is a measure of the degree of magnetization of a material, when it is exposed to a magnetic field. It can be calculated by the equation

$$\chi = M/H$$

Where $M$ is the magnetisation of the material, and $H$ is the applied magnetic field strength.

Aluminium is paramagnetic and gold is diamagnetic, but since they do not possess permanent magnetisation without the presence of an external magnetic field they are considered non-magnetic elements. On the other hand, nickel is a ferromagnetic element, with a large magnetic susceptibility when placed in an external magnetic field. Since the presence of the nickel could cause a significant magnetic field when placed in the 12 T magnet, the magnetic susceptibility for the different kinds of gold-coatings was tested and compared to aluminium.

In order to determine the magnetic susceptibility of each sample, the magnetisation of the material was measured as a function of the applied magnetic field using a vibrating sample magnetometer (VSM) by the department of Physics at the University of Warwick. The results are showed in Figure 6.8.

As expected, the magnetisation increased with the magnetic field: it is lower for the aluminium sample, and increases with the different gold-coatings. The presence of nickel makes a big difference compared to the uncoated and only-gold coated material. These curves were used to calculate the magnetic susceptibility for each sample using the quoted equation. The results are summarised in Table 6.2, which shows the magnetic susceptibility calculated for the aluminium and gold/nickel-coated sample.

These values were compared to the literature for magnetic susceptibility of oxygen-free high conductivity (OFHC) copper ($\chi = 2.46$ x $10^{-5}$ $cm^3$/g) and stainless steel type 316 (SS 316, $\chi = 2.50$ x $10^{-5}$ $cm^3$/g) measured at 12 T [285, 286]. OFHC copper is the material commonly used for the construction of ICR-cells due to its low magnetic

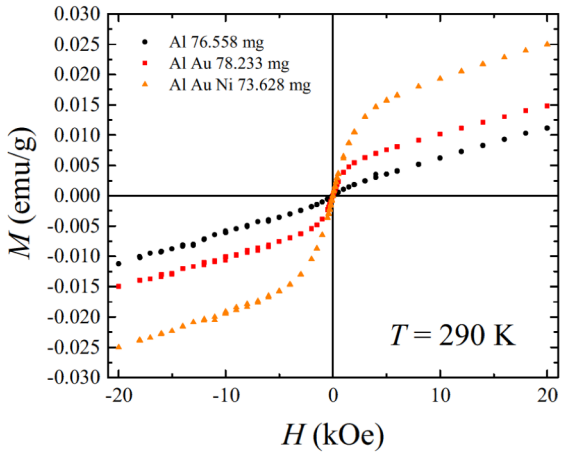



Figure 6.8.: Variation of magnetization of three materials for the construction of the cell with the applied magnetic field. As expected, there is an increase of magnetization with the applied magnetic field H. This is lower for the aluminium sample (black dots), and increases with the gold-coating (red square) and the gold/nickel-coating. Calculations showed that the magnetic susceptibility of the gold/nickel coating is not high enough to influence the magnetic field of the 12 T magnet.

susceptibility. Comparing with the data showed in Table 6.2, it can be seen that the magnetic susceptibility of the gold/nickel-coated sample is of an order of magnitude lower than that of OFHC Cu and SS 316. Although the measurements were done for 1 and 7 T, since $\chi$ diminishes with the magnetic field strength, it is concluded that at 12 T the nickel will have a negligible effect on the magnetic properties of the cell. Another precaution was to measure the change in magnetic susceptibility with temperature. This is of particular importance due to the fact that, in order to improve their performance, ion traps need to be in UHV: to achieve it, the entire part of the instrument under UHV (usually a tube containing part of the transfer optics and the ICR-cell) is periodically subjected to high temperatures in a procedure called *bake-out*, in order for the particles inside it to boil and be pumped out of the system. Figure 6.9 shows that the magnetizations of the aluminium and gold-coated samples decrease with temperature following a linear trend, without significant changes in signal intensities. The gold/nickel-coated sample exhibited higher signals at lower temperatures, which diminished as the temperature increases.

Opposite to what happens in the case of the cell plates, the samples were coated only on one side. These measurements were repeated positioning the gold/nickel coated sample with the coating layer perpendicular and parallel to the applied magnetic field, in order to check whether significant changes in the analysis were present. The curves in Figure 6.10 show no significant differences with Figures 6.8 and 6.9.

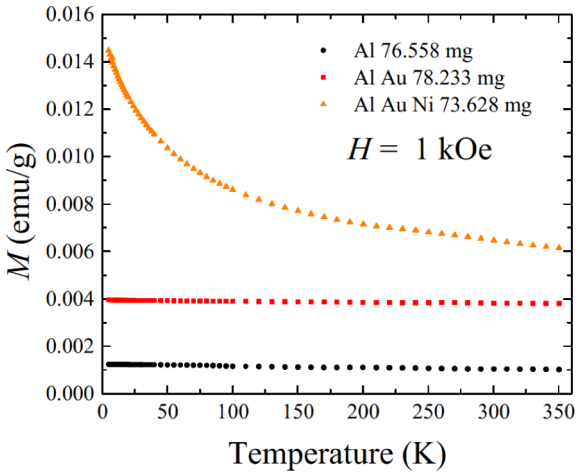



Figure 6.9.: Variation of magnetization with temperature for all the samples at 1 T. The aluminium and gold-coated samples show a linear trend with a small decrease in magnetization with the increase in temperature. The nickel/gold-coated samples exhibits higher values of magnetization at low temperatures, decreasing rapidly with the increase in temperature. The calculated values of $\chi$ for these responses were low enough to consider the material suitable for the cell.

### 6.3.2. Assembly

The cell has been assembled as shown in Figures 6.11 and 6.12. Ceramic rods are connected to the external part of the main electrodes (excitation, detection, inner trapping) in order to give stability to the structure, and hold the sides together by embedding into the outer trapping plates. They have minimal flexibility, in order to support the entire cell structure. The ceramic rods are inserted into the alignment rings to keep a correct alignment of the plates, and are fixed to the plates through aluminium clamps. The clamps are bolted to the plates, but the bolts do not penetrate the inside of the cell, keeping the inner surface completely flat. Ceramic washers are inserted to avoid direct contact between different electrodes, and small ceramic rods (not shown in the picture) are inserted inside the compensation electrodes in order to connect them to their next ones.

The extra two rings outside the cell (towards the background of Figure 6.11) will be used to attach the cell into an existent, although evolving, instrument setup.

The alignment rings were drilled externally, and Teflon plugs were applied to them in order to centre the entire cell structure in the flight tube going into the magnet bore (Figure 6.12). A detailed figure with the shape of the outer alignment ring is available in the Supporting Information (Figure D.1).

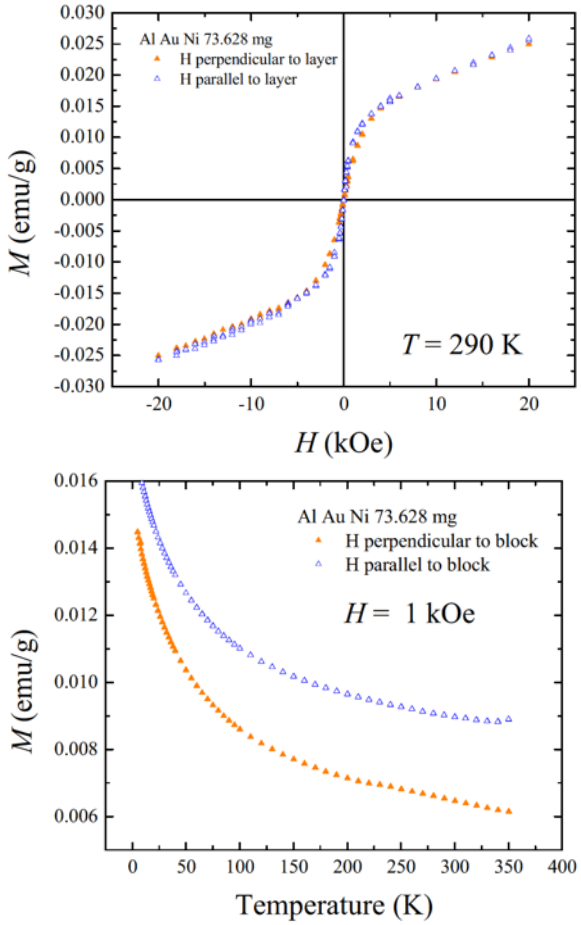



Figure 6.10.: Magnetization variation with the applied magnetic field (left) and temperature (right) for the gold/nickel-coated sample with the coating layer parallel and perpendicular to the applied magnetic field.

Figure 6.11.: The assembled ICR-cell. This picture was taken before the application of the Teflon plugs on the alignment rings. The ICR-cell is connected to two external support rings through four aluminium rods (behind the cell in the picture). The structure allows to isert the cell into a home-built instrument setup.

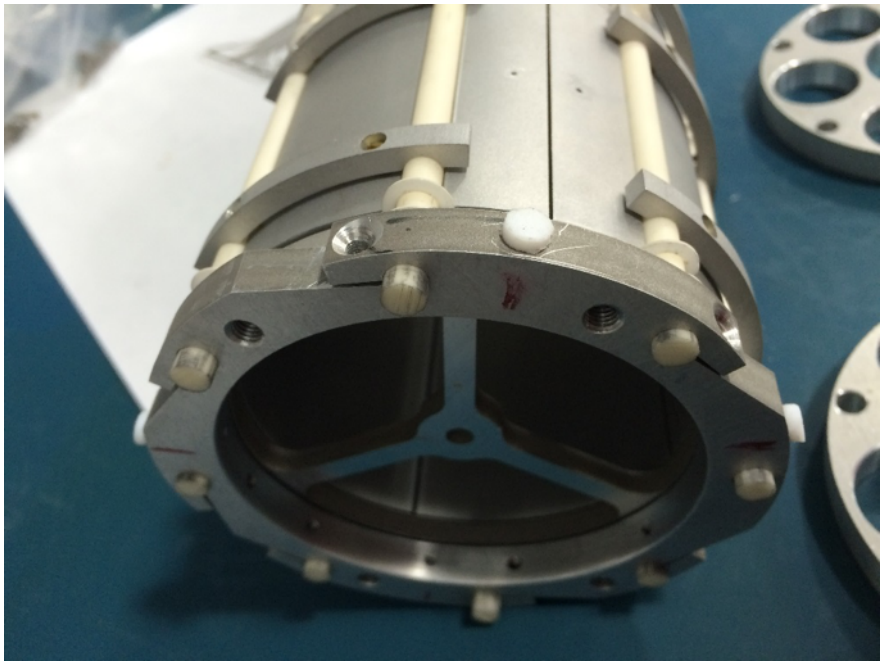


Figure 6.12.: The assembled ICR-cell, with focus on the alignment ring with four Teflon plugs along its external diameter.

### 6.3.3. The Printed Circuit Board (PCB)

Voltages applied to the cell plates will be controlled through the use of a Printed Circuit Board, connected to the ICR-cell with OFHC copper screws. The cell needs a total of two PCBs, put over the gap between two sides containing respectively an excitation plate and a detection plate. A multipin will connect the PCB with the instrument controls. The PCBs have been designed according to the circuit shown in Figure 6.13, and realized by the University of Warwick (Dr Alex Colburn).

Acetate copies of the PCB were generated after the initial measurements, and tested with the screws and the help of Teflon spacers on the assembled cell. The materials used for the PCB need to follow the same precautions applied to the other materials in the cell: the board has been realised in kapton, suitable for Ultra High Vacuum and high temperatures, and OFCH Copper. Figure 6.14 shows a complete PCB for the ICR-cell.

The printed circuit boards were mounted on the ICR-cell, and tested through the use of an oscilloscope in order to confirm their capacitive coupling (Figure D.3 in the Supporting Information).

### 6.3.4. A golden future

After testing the materials, assembling the ICR-cell, the production and tests of the printed circuit boards, and final confirmation that the ion trap worked, the open

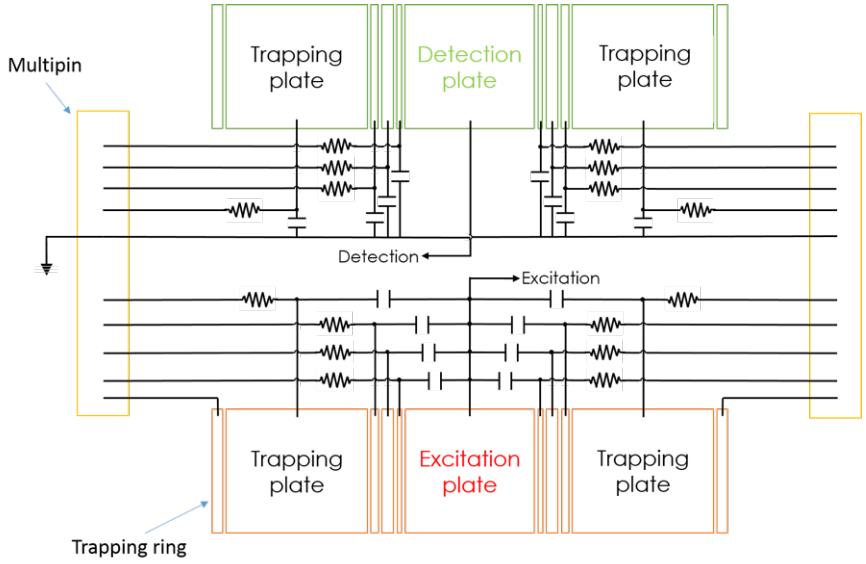



Figure 6.13.: Simple schematic of the circuit required for the PCB. Sides of the cell containing the detection plate and the excitation plate are connected to the same PCB in pairs. The design with real measurements is reported in Figure D.2 in the Supporting Information.

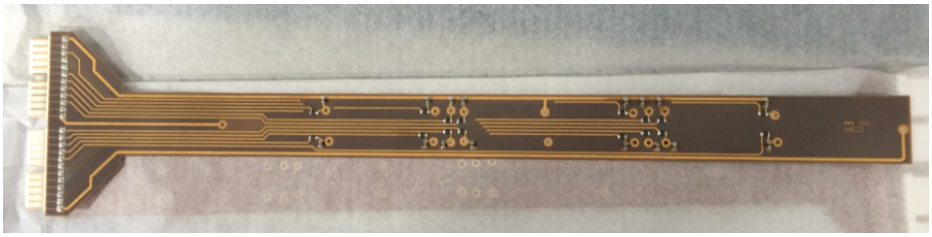


Figure 6.14.: Printed Circuit Board (PCB) realized in kapton and OFHC copper. Small resistors and capacitors can be seen, disposed according to the circuit scheme showed in Figure 6.13.

Figure 6.15.: Ultimated open compensated cylindrical ICR-cell of novel design. All the electrodes are coated with a gold/nickel layer in order to guarantee a smooth consistent, oxidation-resistent surface, necessary to improve ion motion coherence and control. The four cell quadrants are aligned through the use of ceramic rods clamped to each electrode and embedded in outer trapping rings. Ceramic washers allow separation between each plate. The cell is here mounted on a structure made of three aluminium rings and four rods, act to install it in the home-built instrument that the O'Connor group is developing.

cylindrical cell was dismantled in order to be subjected to the gold/nickel coating.

The gold/nickel coating was successful, and the cell was assembled again. At present, the O'Connor laboratory group is developing a home-built FT-ICR mass spectrometer in order to test it, and use it for two-dimensional FT-ICR MS. The improved ion capacity and the capability to longer maintain ion coherence inside the stable confinements of the trap make this novel ICR-cell a promising mean to develop the technique [65, 66]. Figure 6.15 shows the final assembly of the novel open compensated cylindrical cell, which is just a step away from beta-testing. Further detailed images of the assembly are available in the Supporting Information (Figures D.4, D.5, D.6).

# 7. Conclusions

Two-dimensional Fourier transform ion cyclotron resonance mass spectrometry is a viable tandem mass spectrometry technique, which found its applications in the analysis of small molecules and bottom up proteomics.

2D FT-ICR MS was optimised for top-down proteomics, and a comparison study between the technique and standard 1D MS/MS was peformed using CaM. 2D FT-ICR MS gives for both TDP and BUP tandem mass spectrometry information equivalent to multiple 1D experiments targeted at different precursor ions, in a single two-dimensional mass spectrum. The resulting 2D MS experiments allow consistent savings in time and sample consumption. The use of precursor ion scans allows to check at once all the precursors that generated ions of interest, improving the available information about fragmentation pathways. Use of neutral-loss lines allows to obtain information about all the interesting neutral losses at once, and it can be used in BUP studies to assign phosphorylated or glycosylated peptides. The technique was later on optimised for ECD, and the cleavage coverage of CaM was improved of more than 20% by the addition of a single BUP 2D ECD MS experiment.

In order to deepen the sequencing of entire proteins, $MS^n$/2D MS was developed. $MS^n$/2D MS is a technique that unites the high mass accuracy of FT-ICR MS to the simultaneus fragmentation capability of 2D MS, and it produces two-dimensional mass spectra retaining information equivalent to $MS^{n+1}$ of the protein under analysis. $MS^n$/2D MS is performed by isolating and fragmenting a charge state/ion of interest up to $n$-1 times before two-dimensional mass spectrometry. In MS/2D MS, precursor ion scans are given a new use. With the extraction of precursor ion scans it is, in fact, possible to assign all (or many) of the primary fragments with the same terminus on the autocorrelation line derived from the first dissociation.

A 2D FT-ICR MS method for the analysis of polymers was developed, allowing the high resolution tandem mass spectrometry characterisation of a complex mixture of Polysorbate 80, very challenging with standard one-dimensional FT-ICR MS without any sort of separation. Particularly in this study, the use of neutral-loss lines allowed to assign all the linoleated, oleated, and non-esterified species at once. Two-dimensional analysis of the polymeric mixtures (TPGS, Polysorbate 80, and PMMA) revealed the presence of characteristic diagonal lines on the 2D mass spectrum, fingerprints of the different mixtures.

Finally, preliminary studies were performed to develop 2D MS for the analysis of glycans and glycoproteins, and a new ICR-cell was developed to obtain a better performance with 2D MS experiments.

In conclusion, the work of this thesis developed 2D FT-ICR MS, and compared it to standard one-dimensional mass spectrometry, showing the applicability of the technique to top-down proteomics and polymer analysis, and highlighting the several advantages and unique features, such as the use of precursor ion scans and neutral-loss lines. Furthermore, it opened the paths of 2D FT-ICR MS for glycomics and an improved instrumental performance thanks to the use of a more capacious ICR-cell.

# Bibliography


[1] F. Floris and P.B. O'Connor. *Fundamentals and Applications of Fourier Transform Mass Spectrometry*, chapter Fundamentals of 2D-MS, page Submitted. Elsevier, 2018.

[2] Mark P Barrow and Viktor Khristenko. Petroleomics: study of the old and the new. *Biofuels*, 1(5):651–655, sep 2010.

[3] M.B. Comisarow and A.G. Marshall. Fourier transform ion cyclotron resonance spectroscopy. *Chem. Phys. Lett.*, 25(2):282–283, 1974.

[4] Melvin B. Comisarow and Alan G. Marshall. Frequency-sweep fourier transform ion cyclotron resonance spectroscopy. *Chem. Phys. Lett.*, 26(4):489–490, 1974.

[5] Alan G. Marshall, J Tao-chin Lin Wang, and Tom L Riccal. Tailored exctaion for forier transform ion cyclotron resonance mass spectrometry. *J. Am. Chem. Soc.*, 107(26):7893–7897, 1985.

[6] Alan G. Marshall, C L Hendrickson, and G S Jackson. Fourier transform ion cyclotron resonance mass spectrometry: a primer. *Mass Spectrom Rev.*, 17(1):1–35, 1998.

[7] Michael L Easterling and I Jonathan. Amster. Fourier transform mass spectrometry. *J. Mass Spectrom.*, 31:1325–1337, 1996.

[8] A. Gareth Brenton and A. Ruth Godfrey. Accurate mass measurement: Terminology and treatment of data. *J. Am. Soc. Mass. Spectrom.*, 21(11):1821–1835, 2010.

[9] Kermit K. Murray, Robert K. Boyd, Marcos N. Eberlin, G. John Langley, Liang Li, and Yasuhide Naito. Definitions of terms relating to mass spectrometry (IUPAC recommendations 2013). *Pure Appl. Chem.*, 85(7), jan 2013.

[10] Mark P. Barrow, William I Burkitt, and Peter J Derrick. Principles of Fourier transform ion cyclotron resonance mass spectrometry and its application in structural biology. *The Analyst*, 130(i):18–28, 2005.

[11] Xidong Feng, Nigel Clipston, Tracy Brown, Helen Cooper, Uwe Reuther, Andreas Hirsch, Peter J. Derrick, and Thomas Drewello. Generation and detection of [C59N]+ by laser desorption/ionisation high-field Fourier transform

ion cyclotron resonance mass spectrometry. *Rapid Commun. Mass Spectrom.*, 14(5):368–370, mar 2000.

[12] Juan Wei, Oleg N. Antzutkin, Andrei V. Filippov, Dinu Iuga, Pui Yiu Lam, Mark P. Barrow, Ray Dupree, Steven P. Brown, and Peter B. O'Connor. Amyloid Hydrogen Bonding Polymorphism Evaluated by 15N{17O}REAPDOR Solid-State NMR and Ultra-High Resolution Fourier Transform Ion Cyclotron Resonance Mass Spectrometry. *Biochemistry*, 55(14):2065–2068, 2016.

[13] Shu-ping Chen, Melvin B Comisarow, and Receiued I August I. Simple Physical Models for Coulomb-induced Frequency Shifts and Coulomb-induced Inhomogeneous Broadening for Like and Unlike Ions in Fourier Transform Ion Cyclotron Resonance Mass Spectrometry. *Rapid Commun. Mass Spectrom.*, 5(August):450–455, 1991.

[14] Shu-Ping Chen and Melvin B. Comisarow. Modelling coulomb effects in Fourier-transform ion cyclotron resonance mass spectrometry by charged disks and charged cylinders. *Rapid Commun. Mass Spectrom.*, 6(1):1–3, jan 1992.

[15] S Guan, M C Wahl, and A G Marshall. Elimination of frequency drift from Fourier transform ion cyclotron resonance mass spectra by digital quadrature heterodyning: ultrahigh mass resolving power for laser-desorbed molecules. *Anal. Chem.*, 65(24):3647–3653, 1993.

[16] Michael L. Easterling, Todd H. Mize, and I. Jonathan Amster. Routine part-per-million mass accuracy for high-mass ions: Space-charge effects in MALDI FT-ICR. *Anal. Chem.*, 71(3):624–632, 1999.

[17] John T Stults. Minimizing Peak Coalescence: High-Resolution Separation of Isotope Peaks in Partially Deamidated Peptides by Matrix-Assisted Laser Desorption/Ionization {Fourier} Transform Ion Cyclotron Resonance Mass Spectrometry. *Anal Chem*, 69(10):1815–1819, 1997.

[18] P Caravatti and M Allemann. The ' Infinity Cell ': a New Trapped-ion Cell With Radiofrequency Covered Trapping Electrodes for Fourier Transform Ion Cyclotron Resonance Mass Spectrometry. *Org. Mass Spectrom.*, 26:514–518, 1991.

[19] M. S. B. Munson and F. H. Field. Chemical ionization mass spectrometry. i. general introduction. *J. Am. Chem. Soc.*, 88(12):2621–2630, jun 1966.

[20] Andrea Raffaelli and Alessandro Saba. Atmospheric pressure photoionization mass spectrometry. *Mass Spectrom. Rev.*, 22(5):318–331, 2003.

[21] D. B. Robb, T. R. Covey, and a. P. Bruins. Atmospheric pressure photoionization: An ionization method for liquid chromatography - Mass spectrometry. *Anal. Chem.*, 72(15):3653–3659, 2000.

[22] Nadim Hourani and Nikolai Kuhnert. Development of a novel direct-infusion atmospheric pressure chemical ionization mass spectrometry method for the analysis of heavy hydrocarbons in light shredder waste. *Anal. Methods*, 4(3):730, 2012.

[23] Nadim Hourani, Jan T. Andersson, Isabelle Möller, Ma'an Amad, Matthias Witt, and S. Mani Sarathy. Atmospheric pressure chemical ionization fourier transform ion cyclotron resonance mass spectrometry for complex thiophenic mixture analysis. *Rapid Commun. Mass Spectrom.*, 27(21):2432–2438, oct 2013.

[24] Clotilde Le Vot, Carlos Afonso, Claude Beaugrand, and Jean-Claude Tabet. Implementation of a penning ionization source on a FTICR instrument with ion funnel optics. *Int. J. Mass Spectrom.*, 306(2-3):150–158, sep 2011.

[25] Michael. Karas and Franz. Hillenkamp. Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons. *Anal. Chem.*, 60(20):2299–2301, 1988.

[26] Victor V. Laiko, Michael A. Baldwin, and Alma L. Burlingame. Atmospheric pressure matrix-assisted laser desorption/ionization mass spectrometry. *Anal. Chem.*, 72(4):652–657, 2000.

[27] Bogdan A Budnik, Susanne C Moyer, Jason L Pittman, Vera B Ivleva, Ulf Sommer, Catherine E Costello, and Peter B O'Connor. High pressure MALDI-FTMS: implications for proteomics. *Int. J. Mass Spectrom.*, 234(1-3):203–212, may 2004.

[28] Peter B. O'Connor, Bogdan A. Budnik, Vera B. Ivleva, Parminder Kaur, Susanne C. Moyer, Jason L. Pittman, and Catherine E. Costello. A high pressure matrix-assisted laser desorption ion source for Fourier transform mass spectrometry designed to accommodate large targets with diverse surfaces. *J. Am. Soc. Mass. Spectrom.*, 15(1):128–132, jan 2004.

[29] Masamichi Yamashita and John B. Fenn. Negative ion production with the electrospray ion source. *J. Phys. Chem.*, 88(20):4671–4675, sep 1984.

[30] Masamichi Yamashita and John B. Fenn. Electrospray ion source. another variation on the free-jet theme. *J. Phys. Chem.*, 88(125):4451–4459, sep 1984.

[31] J B Fenn, M Mann, C K Meng, S F Wong, and C M Whitehouse. Electrospray ionization for mass spectrometry of large biomolecules. *Science*, 246(4926):64 LP – 71, oct 1989.

[32] Matthias Wilm. Principles of electrospray ionization. *Mol. Cell. Proteomics*, 10(7):M111.009407, 2011.

[33] C S Ho, C W K Lam, M H M Chan, R C K Cheung, L K Law, L C W Lit, K F

Ng, M W M Suen, and H L Tai. Electrospray ionisation mass spectrometry: principles and clinical applications. *Clin. Biochem.*, 24(1):3–12, 2003.

[34] J. V. Iribarne. On the evaporation of small ions from charged droplets. *The Journal of Chemical Physics*, 64(6):2287, 1976.

[35] Paul Kebarle and Udo H. Verkerk. Electrospray: From ions in solution to ions in the gas phase, what we know now. *Mass Spectrom. Rev.*, 28(6):898–917, nov 2009.

[36] Anthony T. Iavarone and Evan R. Williams. Mechanism of charging and supercharging molecules in electrospray ionization. *J. Am. Chem. Soc.*, 125(8):2319–2327, feb 2003.

[37] Malcolm Dole, L. L. Mack, R. L. Hines, R. C. Mobley, L. D. Ferguson, and M. B. Alice. Molecular Beams of Macroions. *J. Chem. Phys*, 49(5):2240–2249, sep 1968.

[38] Lars Konermann, Elias Ahadi, Antony D. Rodriguez, and Siavash Vahidi. Unraveling the mechanism of electrospray ionization. *Anal. Chem.*, 85(1):2–9, nov 2013.

[39] Jeonghoon Lee and Peter T. A. Reilly. Limitation of time-of-flight resolution in the ultra high mass range. *Anal. Chem.*, 83(15):5831–5833, aug 2011.

[40] Jeonghoon Lee, Huijuan Chen, Tiancheng Liu, Clifford E. Berkman, and Peter T. A. Reilly. High resolution time-of-flight mass analysis of the entire range of intact singly-charged proteins. *Anal. Chem.*, 83(24):9406–9412, dec 2011.

[41] Ruedi Aebersold and Matthias Mann. Mass spectrometry-based proteomics. *Nature*, 422(6928):198–207, 2003.

[42] Roman Korner, Matthias Wilrn, Kenneth Morand, Michael Schubert, and Matthias Mann. Nano Electrospray Combined with a Quadrupole Ion Trap for the Analysis of Peptides and Protein Digests *. *J. Am. Soc. Mass. Spectrom.*, 0305(95):150–156, feb 1996.

[43] Matthias Wilm and Matthias Mann. Analytical Properties of the Nanoelectrospray Ion Source. *Anal. Chem.*, 68(1):1–8, jan 1996.

[44] Reinhard Juraschek, Thomas DÃŒlcks, and Michael Karas. Nanoelectrospray—more than just a minimized-flow electrospray ionization source. *J. Am. Soc. Mass. Spectrom.*, 10(4):300–308, apr 1999.

[45] Andrea Schmidt, Michael Karas, and Thomas DÃŒlcks. Effect of different solution flow rates on analyte ion signals in nano-ESI MS, or: when does ESI turn into nano-ESI? *J. Am. Soc. Mass. Spectrom.*, 14(5):492–500, may 2003.

[46] Ayman El-Faramawy, K. W. Michael Siu, and Bruce A. Thomson. Efficiency of nano-electrospray ionization. *J. Am. Soc. Mass. Spectrom.*, 16(10):1702–1707, oct 2005.

[47] Mridul Kanti Mandal, Kentaro Yoshimura, Subhrakanti Saha, Satoshi Ninomiya, Md. Obaidur Rahman, Zhan Yu, Lee Chuin Chen, Yasuo Shida, Sen Takeda, Hiroshi Nonami, and Kenzo Hiraoka. Solid probe assisted nanoelectrospray ionization mass spectrometry for biological tissue diagnostics. *Analyst*, 137(20):4658, 2012.

[48] Rebecca A. Jockusch, Paul D. Schnier, William D. Price, Eric. F. Strittmatter, Plamen A. Demirev, and Evan R. Williams. Effects of charge state on fragmentation pathways, dynamics, and activation energies of ubiquitin ions measured by blackbody infrared radiative dissociation. *Analytical Chemistry*, 69(6):1119–1126, mar 1997.

[49] Albert J R Heck. Native mass spectrometry: a bridge between interactomics and structural biology. *Nat. Methods*, 5(11):927–933, nov 2008.

[50] Sheng Yin and Joseph A Loo. International Journal of Mass Spectrometry Top-down mass spectrometry of supercharged native protein â ligand complexes. *Int. J. Mass Spectrom.*, 300(2-3):118–122, 2011.

[51] Michael T Marty, Hao Zhang, Weidong Cui, Robert E Blankenship, Michael L Gross, and Stephen G Sligar. Native Mass Spectrometry Characterization of Intact Nanodisc Lipoprotein Complexes. *Anal. Chem.*, 84(21):8–11, oct 2012.

[52] Huilin Li, Jeremy J Wolff, Steve L Van Orden, and Joseph A Loo. Native top-down electrospray ionization-mass spectrometry of 158 kDa protein complex by high-resolution fourier transform ion cyclotron resonance mass spectrometry. *Anal. Chem.*, 86(1), 2014.

[53] Ryan T. Kelly, Aleksey V. Tolmachev, Jason S. Page, Keqi Tang, and Richard D. Smith. The ion funnel: Theory, implementations, and applications. *Mass Spectrom. Rev.*, 47(12):n/a–n/a, 2009.

[54] Fred W. McLafferty, Kathrin Breuker, Mi Jin, Xuemei Han, Giuseppe Infusini, Honghai Jiang, Xianglei Kong, and Tadhg P. Begley. Top-down MS, a powerful complement to the high capabilities of proteolysis proteomics. *FEBS J.*, 274(24):6256–6268, nov 2007.

[55] Peter B O'connor, Jason L Pittman, Bruce a Thomson, Bogdan a Budnik, Jason C Cournoyer, Judith Jebanathirajah, Cheng Lin, Susanne Moyer, and Cheng Zhao. A new hybrid electrospray Fourier transform mass spectrometer: design and performance characteristics. *Rapid Commun. Mass Spectrom.*, 20(2):259–66, 2006.

[56] Jennifer S. Cobb, Michael L. Easterling, and Jeffrey N. Agar. Structural characterization of intact proteins is enhanced by prevalent fragmentation pathways rarely observed for peptides. *J. Am. Soc. Mass. Spectrom.*, 21(6):949–959, 2010.

[57] Donald J. Douglas, Aaron J. Frank, and Dunmin Mao. Linear ion traps in mass spectrometry. *Mass Spectrom. Rev.*, 24(1):1–29, 2005.

[58] Alan G. Marshall. Milestones in fourier transform ion cyclotron resonance mass spectrometry technique development. *Int. J. Mass Spectrom.*, 200(1-3):331–356, dec 2000.

[59] J.W. Gauthier, T.R. Trautman, and D.B. Jacobson. Sustained off-resonance irradiation for collision-activated dissociation involving Fourier transform mass spectrometry. Collision-activated dissociation technique that emulates infrared multiphoton dissociation. *Anal. Chim. Acta*, 246(1):211–225, may 1991.

[60] Eugene N Nikolaev, Yury I Kostyukevich, and Gleb N Vladimirov. FOURIER TRANSFORM ION CYCLOTRON RESONANCE ( FT ICR ) MASS SPECTROMETRY : THEORY AND SIMULATIONS. *Mass Spectrom. Rev.*, 35(2):219–258, feb 2014.

[61] Steven C. Beu and David A. Laude. Open trapped ion cell geometries for Fourier transform ion cyclotron resonance mass spectrometry. *International Journal of Mass Spectrometry and Ion Processes*, 112(2-3):215–230, jan 1992.

[62] Steven C Beu and David A Laude. Elimination of Axial Ejection during Excitation with a Capacitively Coupled Open Trapped- Ion Cell for Fourier Transform Ion Cyclotron Resonance Mass Spectrometry. *Anal. Chem.*, 180(12):177–180, 1992.

[63] Aleksey V Tolmachev, Errol W Robinson, Si Wu, Hyuk Kang, Natacha M Lourette, Ljiljana Pasa-Tolić, and Richard D Smith. Trapped-ion cell with improved DC potential harmonicity for FT-ICR MS. *J. Am. Soc. Mass. Spectrom.*, 19(4):586–97, apr 2008.

[64] Adam M Brustkern, Don L Rempel, and Michael L Gross. An Electrically Compensated Trap Designed to Eighth Order for FT-ICR Mass Spectrometry. *J. Am. Soc. Mass. Spectrom.*, 2008.

[65] Adam M. Brustkern, Don L. Rempel, and Michael L. Gross. A Tuning Method for Electrically Compensated Ion Cyclotron Resonance Mass Spectrometer Traps. *J. Am. Soc. Mass. Spectrom.*, 21(3):451–454, 2010.

[66] Adam M Brustkern, Don L Rempel, and Michael L Gross. Ion Behavior in an Electrically Compensated Ion Cyclotron Resonance Trap. *Int. J. Mass Spectrom.*, 300(2-3):143–148, mar 2011.

[67] Ivan a. Boldin and Eugene N. Nikolaev. Fourier transform ion cyclotron resonance cell with dynamic harmonization of the electric field in the whole volume by shaping of the excitation and detection electrode assembly. *Rapid Commun. Mass Spectrom.*, 25(July 2010):122–126, 2011.

[68] Eugene N Nikolaev, Ivan a Boldin, Roland Jertz, and Gökhan Baykut. Initial experimental characterization of a new ultra-high resolution FTICR cell with dynamic harmonization. *J. Am. Soc. Mass. Spectrom.*, 22(7):1125–33, jul 2011.

[69] Eugene N Nikolaev, Roland Jertz, and Anton Grigoryev. Fine Structure in Isotopic Peak Distributions Measured Using a Dynamically Harmonized Fourier Transform Ion Cyclotron Resonance Cell at 7 T. *Anal. Chem.*, 84(5):2275–2283, feb 2012.

[70] Yulin Qi, Matthias Witt, Roland Jertz, Gökhan Baykut, Mark P. Barrow, Eugene N Nikolaev, and Peter B O'Connor. Absorption-mode spectra on the dynamically harmonized Fourier transform ion cyclotron resonance cell. *Rapid Commun. Mass Spectrom.*, 26(17):2021–6, sep 2012.

[71] Konstantin O. Nagornov, Anton N. Kozhinov, Oleg Y. Tsybin, and Yury O. Tsybin. Ion trap with narrow aperture detection electrodes for fourier transform ion cyclotron resonance mass spectrometry. *J. Am. Soc. Mass. Spectrom.*, 26(5):741–751, 2015.

[72] Franklin E. Leach, Randolph Norheim, Gordon Anderson, and Ljiljana Pasa-Tolic. Application of Printed Circuit Board Technology to FT-ICR MS Analyzer Cell Construction and Prototyping. *J. Am. Soc. Mass Spectrom.*, 25(12):2069–2072, jul 2014.

[73] Sung-Gun Park, Gordon A. Anderson, Arti T. Navare, and James E. Bruce. Parallel spectral acquisition with an ion cyclotron resonance cell array. *Anal. Chem.*, 88(2):1162–1168, dec 2015.

[74] Konstantin Aizikov, Raman Mathur, and Peter B. O'Connor. The Spontaneous Loss of Coherence Catastrophe in Fourier Transform Ion Cyclotron Resonance Mass Spectrometry. *J. Am. Soc. Mass. Spectrom.*, 20(2):247–256, 2009.

[75] Alan G Marshall, Christopher L Hendrickson, and Stone D.-H. Shi. Scaling MS plateaus with high-resolution FT-ICRMS. *Analytical Chemistry*, 74(9):252A–259A, 2002.

[76] M P Barrow, M Witt, J V Headly, and K M Peru. Athabasca Oil Sands Process Water Characterization by Atmospheric Pressure Photoionization and Electrospray Ionization Fourier Transform Ion Cyclotron Resonance Mass Spectrometry. *Anal. Chem.*, 82(9):3727–3735, 2010.

[77] Shenheng Guan and Alan G. Marshall. Stored waveform inverse Fourier transform (SWIFT) ion excitation in trapped-ion mass spectometry: Theory and applications. *Int. J. Mass Spectrom. Ion Processes*, 157-158:5–37, 1996.

[78] Malgorzata A. Kaczorowska, Anna C G Hotze, Michael J. Hannon, and Helen J. Cooper. Electron Capture Dissociation Mass Spectrometry of Metallo-Supramolecular Complexes. *J. Am. Soc. Mass. Spectrom.*, 21(2):300–309, 2010.

[79] Alan G. Marshall and Christopher L. Hendrickson. Fourier transform ion cyclotron resonance detection: principles and experimental configurations. *Int. J. Mass Spectrom.*, 215(1-3):59–75, apr 2002.

[80] L J de Koning, N M M Nibbering, S L van Orden, and F H Laukien. Mass selection of ions in a Fourier transform ion cyclotron resonance trap using correlated harmonic excitation fields (CHEF). *Int. J. Mass Spectrom.*, 165:209–219, 1997.

[81] Bogdan Bogdanov and Richard D. Smith. Proteomics by fticr mass spectrometry: TOP down and bottom up. *Mass Spectrom. Rev.*, 24:168–200, 2005.

[82] Brian T Chait. Mass Spectrometry : Bottom-Up or Top-Down? *Science*, 314:65–66, 2006.

[83] Andrew J. Forbes, Matthew T. Mazur, Neil L. Kelleher, Hiten M. Patel, and Christopher T. Walsh. Toward efficient analysis of 70 kDa proteins with 100% sequence coverage. *Eur. J. Mass Spectrom.*, 7(2):81–87, apr 2001.

[84] Cong Wu, John C Tran, Leonid Zamdborg, Kenneth R Durbin, Mingxi Li, Dorothy R Ahlf, Bryan P Early, Paul M Thomas, Jonathan V Sweedler, and Neil L Kelleher. A protease for 'middle-down' proteomics. *Nat. Methods*, 9(8):822–824, jun 2012.

[85] P. Roepstorff and J. Fohlman. Letter to the editors. *Biol. Mass. Spectrom.*, 11(11):601–601, nov 1984.

[86] Bruno Domon and Catherine E. Costello. A systematic nomenclature for carbohydrate fragmentations in FAB-MS/MS spectra of glycoconjugates. *Glycoconjugate J.*, 5(4):397–409, 1988.

[87] Joseph A. Loo, Harold R. Udseth, Richard D. Smith, and J. H. Futrell. Collisional effects on the charge distribution of ions from large molecules, formed by electrospray-ionization mass spectrometry. *Rapid Commun. Mass Spectrom.*, 2(10):207–210, dec 1988.

[88] Scott A. McLuckey. Principles of collisional activation in analytical mass spectrometry. *J. Am. Soc. Mass. Spectrom.*, 3(6):599–614, 1992.

[89] Béla Paizs and Sándor Suhai. Fragmentation pathways of protonated peptides. *Mass Spectrom. Rev.*, 24(4):508–48, 2005.

[90] J. Mitchell Wells and Scott A. McLuckey. CollisionâInduced Dissociation (CID) of Peptides and Proteins. In *Methods in Enzymology*, pages 148–185. Elsevier, 2005.

[91] Ashok R. Dongré, Jennifer L. Jones, Árpád Somogyi, and Vicki H. Wysocki. Influence of peptide composition, gas-phase basicity, and chemical modification on fragmentation efficiency: evidence for the mobile proton model. *J. Am. Chem. Soc.*, 118(35):8365–8374, jan 1996.

[92] Vicki H. Wysocki, George Tsaprailis, Lori L. Smith, and Linda A. Breci. Mobile and localized protons: a framework for understanding peptide dissociation. *J. Mass Spectrom.*, 35(12):1399–1406, dec 2000.

[93] R Zubarev, N L Kelleher, and F W McLafferty. Electron capture dissociation of multiply charged protein cations. A Nonergodic process. *J. Am. Chem. Soc*, 120(16):3265–3266, 1998.

[94] Roman A. Zubarev, David M. Horn, Einar K. Fridriksson, Neil L. Kelleher, Nathan A. Kruger, Mark A. Lewis, Barry K. Carpenter, and Fred W. McLafferty. Electron Capture Dissociation for Structural Characterization of Multiply Charged Protein Cations. *Anal. Chem.*, 72(3):563–573, feb 2000.

[95] F W McLafferty, David M Horn, Kathrin Breuker, Ying Ge, Mark A Lewis, Blas Cerda, Roman A Zubarev, and Barry K Carpenter. Electron capture dissociation of gaseous multiply charged ions by Fourier-transform ion cyclotron resonance. *J. Am. Soc. Mass. Spectrom.*, 12(3):245–9, 2001.

[96] John E P Syka, Joshua J Coon, Melanie J Schroeder, Jeffrey Shabanowitz, and Donald F Hunt. Peptide and protein sequence analysis by electron transfer dissociation mass spectrometry. *Proc. Natl. Acad. Sci. U.S.A.*, 101(26):9528–33, jun 2004.

[97] Helen J Cooper, Kristina Håkansson, and Alan G Marshall. The role of electron capture dissociation in biomolecular analysis. *Mass Spectrom. Rev.*, 24(2):201–222, 2005.

[98] Konstantin O Zhurov, Luca Fornelli, Matthew D Wodrich, Ünige a Laskay, and Yury O Tsybin. Principles of electron capture and transfer dissociation mass spectrometry applied to peptide and protein structure analysis. *Chem. Soc. Rev.*, 42(12):5014–30, jun 2013.

[99] Roman a Zubarev. Electron-capture dissociation tandem mass spectrometry. *Curr. Opin. Biotechnol.*, 15(1):12–6, feb 2004.

[100] Julia Wiesner, Thomas Premsler, and Albert Sickmann. Application of electron transfer dissociation (ETD) for the analysis of posttranslational modifications. *Proteomics*, 8(21):4466–4483, nov 2008.

[101] Min-Sik Kim and Akhilesh Pandey. Electron transfer dissociation mass spectrometry in proteomics. *Proteomics*, 12(4-5):530–42, 2012.

[102] Mirela Sarbu, Roxana M. Ghiulai, and Alina D. Zamfir. Recent developments and applications of electron transfer dissociation mass spectrometry in proteomics. *Amino Acids*, 46(7):1625–1634, 2014.

[103] František TurecĬek. N-C$\alpha$ Bond Dissociation Energies and Kinetics in Amide and Peptide Radicals. Is the Dissociation a Non-ergodic Process? *J. Am. Chem. Soc.*, 125(19):5954–5963, may 2003.

[104] Nancy Leymarie, Catherine E Costello, and Peter B. O'Connor. Electron Capture Dissociation Initiates a Free Radical Reaction Cascade. *J. Am. Chem. Soc.*, 125(29):8949–8958, 2003.

[105] Jack Simons. Mechanisms for SâS and NâC$\alpha$ bond cleavage in peptide ECD and ETD mass spectrometry. *Chem. Phys. Lett.*, 484(4-6):81–95, jan 2010.

[106] Andrea F. Lopez-Clavijo. *Tandem mass spectrometry of non-enzymatically glycated peptides and proteins.* PhD thesis, University of Warwick, October 2013.

[107] Frank Kjeldsen, Kim F. Haselmann, Bogdan A. Budnik, Frank Jensen, and Roman A. Zubarev. Dissociative capture of hot (3-13 eV) electrons by polypeptide polycations: An efficient process accompanied by secondary fragmentation. *Chem. Phys. Lett.*, 356(3-4):201–206, 2002.

[108] J. J. Coon, B. Ueberheide, J. E. P. Syka, D. D. Dryhurst, J. Ausio, J. Shabanowitz, and D. F. Hunt. Protein identification using sequential ion/ion reactions and tandem mass spectrometry. *Proc. Natl. Acad. Sci. U.S.A.*, 102(27):9463–9468, jun 2005.

[109] Y. M Eva Fung, Christopher M. Adams, and Roman A. Zubarev. Electron ionization dissociation of singly and multiply charged peptides. *J. Am. Chem. Soc.*, 2009.

[110] J Wolff, T Laremore, H Aslam, R Linhardt, and I Amster. Electron-induced dissociation of glycosaminoglycan tetrasaccharides. *J. Am. Soc. Mass. Spectrom.*, 19(10):1449–1458, oct 2008.

[111] Rebecca H. Wills and Peter B. O'Connor. Structural characterization of actinomycin D using multiple ion isolation and electron induced dissociation. *J. Am. Soc. Mass. Spectrom.*, 25(2):186–195, 2014.

[112] Hadi Lioe and Richard A. J. O'Hair. Comparison of collision-induced dissociation and electron-induced dissociation of singly protonated aromatic amino acids, cystine and related simple peptides using a hybrid linear ion trap–FT-ICR mass spectrometer. *Anal. Bioanal.Chem.*, 389(5):1429–1437, sep 2007.

[113] Jackie A. Mosely, Michael J. P. Smith, Aruna S. Prakash, Martin Sims, and Anthony W. T. Bristow. Electron-induced dissociation of singly charged organic cations as a tool for structural characterization of pharmaceutical type molecules. *Anal. Chem.*, 83(11):4068–4075, jun 2011.

[114] Aruna S Prakash, Michael J P Smith, Zied Kaabia, Glenn Hurst, Ci Yan, Martin Sims, Anthony W T Bristow, Peter Stokes, David Parker, and Jackie A Mosely. Using Electron Induced Dissociation (EID) on an LC Time-Scale to Characterize a Mixture of Analogous Small Organic Molecules. *J. Am. Soc. Mass Spectrom*, 23:850–857, 2012.

[115] Juan Wei, Huilin Li, Mark P Barrow, and Peter B O'Connor. Structural characterization of chlorophyll-a by high resolution tandem mass spectrometry. *J. Am. Soc. Mass. Spectrom.*, 24(5):753–760, 2013.

[116] Roman A. Zubarev, Alexander R. Zubarev, and Mikhail M. Savitski. Electron capture/transfer versus collisionally activated/induced dissociations: Solo or duet? *J. Am. Soc. Mass. Spectrom.*, 19(6):753–761, jun 2008.

[117] Yuping Zhou, Jia Dong, and Richard W. Vachet. Electron transfer dissociation of modified peptides and proteins. *Current Pharm. Biotechnol.*, 12(10):1558–1567, oct 2011.

[118] Yuan Mao, Santosh G. Valeja, Jason C. Rouse, Christopher L. Hendrickson, and Alan G. Marshall. Top-down structural analysis of an intact monoclonal antibody by electron capture dissociation-fourier transform ion cyclotron resonance-mass spectrometry. *Analytical Chemistry*, 85(9):4239–4246, apr 2013.

[119] Sandrine Bourgoin-Voillard, Nancy Leymarie, and Catherine E. Costello. Top-down tandem mass spectrometry on RNase A and B using a Qh/FT-ICR hybrid mass spectrometer. *Proteomics*, 14(10):1174–1184, 2014.

[120] D. M. Horn, Y. Ge, and F. W. McLafferty. Activated ion electron capture dissociation for mass spectral sequencing of larger (42 kDa) proteins. *Anal. Chem.*, 72(20):4778–4784, 2000.

[121] H. Oh and F. W. Mclafferty. A varieet of activation methods employed in 'Activated-ion' electron capture. *Bull. Korean. Chem. Soc.*, 27(3):389–394, 2006.

[122] Sharon J. Pitteri, Paul A. Chrisman, and Scott A. McLuckey. Electron-transfer ion/ion reactions of doubly protonated peptides: Effect of elevated bath gas temperature. *Anal. Chem.*, 77(17):5662–5669, 2005.

[123] Aaron R. Ledvina, Graeme C. McAlister, Myles W. Gardner, Suncerae I. Smith, James A. Madsen, Jae C. Schwartz, George C. Stafford, John E P Syka, Jennifer S. Brodbelt, and Joshua J. Coon. Infrared photoactivation reduces peptide folding and hydrogenatom migration following ETD tandem mass spectrometry. *Angew. Chem. Int. Ed*, 48(45):8526–8528, 2009.

[124] Danielle L. Swaney, Graeme C. McAlister, Matthew Wirtala, Jae C. Schwartz, John E P Syka, and Joshua J. Coon. Supplemental activation method for high-efficiency electron-transfer dissociation of doubly protonated peptide precursors. *Anal. Chem.*, 79(2):477–485, 2007.

[125] Victor A. Mikhailov and Helen J. Cooper. Activated ion electron capture dissociation (AI ECD) of proteins: Synchronization of infrared and electron irradia-

tion with ion magnetron motion. *J. Am. Soc. Mass. Spectrom.*, 20(5):763–771, may 2009.

[126] Bogdan A. Budnik, Kim F. Haselmann, and Roman A. Zubarev. Electron detachment dissociation of peptide di-anions : an electron-hole recombination phenomenon. *Chem. Phys. Lett.*, 342(July):299–302, 2001.

[127] Frank Kjeldsen, Oleg A. Silivra, Igor A. Ivonin, Kim F. Haselmann, Mikhail Gorshkov, and Roman A. Zubarev. C$\alpha$-C backbone fragmentation dominates in electron detachment dissociation of gas-phase polypeptide polyanions. *Chem. Eur. J.*, 11(6):1803–1812, 2005.

[128] Frank Kjeldsen, Ole B. Hørning, Søren S. Jensen, Anders M B Giessing, and Ole N. Jensen. Towards Liquid Chromatography Time-Scale Peptide Sequencing and Characterization of Post-Translational Modifications in the Negative-Ion Mode Using Electron Detachment Dissociation Tandem Mass Spectrometry. *J. Am. Soc. Mass. Spectrom.*, 19(8):1156–1162, 2008.

[129] Hangtian Song and Kristina Håkansson. Electron detachment dissociation and negative ion infrared multiphoton dissociation of electrosprayed intact proteins. *Anal. Chem.*, 84(2):871–876, 2012.

[130] Anastasia Kalli, Gabriela Grigorean, and Kristina Håkansson. Electron induced dissociation of singly deprotonated peptides. *J. Am. Soc. Mass. Spectrom.*, 22(12):2209–2221, 2011.

[131] Hyun Ju Yoo, Ning Wang, Shuyi Zhuang, Hangtian Song, and Kristina Håkansson. Negative-ion electron capture dissociation: Radical-driven fragmentation of charge-increased gaseous peptide anions. *J. Am. Chem. Soc.*, 133(42):16790–16793, 2011.

[132] Katherine E. Hersberger and Kristina Håkansson. Characterization of O-sulfopeptides by negative ion mode tandem mass spectrometry: Superior performance of negative ion electron capture dissociation. *Anal. Chem.*, 84(15):6370–6377, 2012.

[133] Joshua J. Coon, Jeffrey Shabanowitz, Donald F. Hunt, and J. E P Syka. Electron transfer dissociation of peptide anions. *J. Am. Soc. Mass. Spectrom.*, 16(6):880–882, 2005.

[134] Malwina Huzarska, Israel Ugalde, Desmond A. Kaplan, Ralf Hartmer, Michael L. Easterling, and Nick C. Polfer. Negative electron transfer dissociation of deprotonated phosphopeptide anions: choice of radical cation reagent and competition between electron and proton transfer. *Anal. Chem.*, 82(7):2873–2878, 2010.

[135] Neil G. Rumachik, Graeme C. McAlister, Jason D. Russell, Derek J. Bailey, Craig D. Wenger, and Joshua J. Coon. Characterizing peptide neutral losses

induced by negative electron-transfer dissociation (NETD). *J. Am. Soc. Mass. Spectrom.*, 23(4):718–727, 2012.

[136] Graeme C. McAlister, Jason D. Russell, Neil G. Rumachik, Alexander S. Hebert, John E P Syka, Lewis Y. Geer, Michael S. Westphall, David J. Pagliarini, and Joshua J. Coon. Analysis of the acidic proteome with negative electron-transfer dissociation mass spectrometry. *Anal. Chem.*, 84(6):2875–2882, 2012.

[137] Jared B. Shaw, Desmond A. Kaplan, and Jennifer S. Brodbelt. Activated ion negative electron transfer dissociation of multiply charged peptide anions. *Anal. Chem.*, 85(9):4721–4728, 2013.

[138] Nicholas M. Riley, Matthew J. P. Rush, Christopher M. Rose, Alicia L. Richards, Nicholas W. Kwiecien, Derek J. Bailey, Alexander S. Hebert, Michael S. Westphall, and Joshua J. Coon. The Negative Mode Proteome with Activated Ion Negative Electron Transfer Dissociation. *Mol. Cell. Proteomics*, 14:2644–2660, 2015.

[139] James P. Reilly. Ultraviolet photofragmentation of biomolecular ions. *Mass Spectrom. Rev.*, 28(3):425–447, may 2009.

[140] Tony Ly and Ryan R. Julian. Ultraviolet photodissociation: developments towards applications for mass-spectrometry-based proteomics. *Angewandte Chemie - International Edition*, 48(39):7130–7137, 2009.

[141] Jennifer S Brodbelt. Photodissociation mass spectrometry: new tools for characterization of biological molecules. *Chem. Soc. Rev.*, 43(8):2757–83, apr 2014.

[142] Daniel P. Little, J. Paul Speir, Michael W. Senko, Peter B. O'Connor, and Fred W. McLafferty. Infrared multiphoton dissociation of large multiply charged ions for biomolecule sequencing. *Anal. Chem.*, 66(18):2809–2815, 1994.

[143] Jennifer S. Brodbelt and Jeffrey J. Wilson. Infrared multiphoton dissociation in quadrupole ion traps. *Mass Spectrom. Rev.*, 28(3):390–424, may 2009.

[144] Myles W Gardner, Suncerae I Smith, Aaron R Ledvina, James A Madsen, Joshua J Coon, Jae C Schwartz, George C Stafford, and Jennifer S Brodbelt. Infrared Multiphoton Dissociation of Peptide Cations in a Dual Pressure Linear Ion Trap Mass Spectrometer. *Anal. Chem.*, 81(19):8109–8118, oct 2009.

[145] James a Madsen, Myles W Gardner, Suncerae I Smith, Aaron R Ledvina, Joshua J Coon, Jae C Schwartz, George C. Stafford, and Jennifer S Brodbelt. Top-Down Protein Fragmentation by Infrared Multiphoton Dissociation in a Dual Pressure Linear Ion Trap. *Anal. Chem.*, 81(21):8677–8686, nov 2009.

[146] Aaron R. Ledvina, M. Violet Lee, Graeme C. McAlister, Michael S. Westphall, and Joshua J. Coon. Infrared multiphoton dissociation for quantitative shotgun proteomics. *Anal. Chem.*, 84(10):4513–4519, 2012.

[147] Matthew S. Thompson, Weidong Cui, and James P. Reilly. Fragmentation of Singly Charged Peptide Ions by Photodissociation at $\lambda$=157 nm. *Angew. Chem. Int. Ed.*, 28(3):425–447, may 2004.

[148] James A. Madsen, Daniel R. Boutz, and Jennifer S. Brodbelt. Ultrafast ultraviolet photodissociation at 193 nm and its applicability to proteomic workflows. *J. Proteome Res.*, 9(8):4205–4214, 2010.

[149] Weidong Cui, Matthew S. Thompson, and James P. Reilly. Pathways of peptide ion fragmentation induced by vacuum ultraviolet light. *J. Am. Soc. Mass. Spectrom.*, 16(8):1384–1398, 2005.

[150] Matthew S. Thompson, Weidong Cui, and James P. Reilly. Factors That Impact the Vacuum Ultraviolet Photofragmentation of Peptide Ions. *J. Am. Soc. Mass. Spectrom.*, 18(8):1439–1452, 2007.

[151] Ramakrishnan Parthasarathi, Yi He, James P. Reilly, and Krishnan Raghavachari. New insights into the vacuum UV photodissociation of peptides. *J. Am. Chem. Soc.*, 132(5):1606–1610, 2010.

[152] Yi He, Nathaniel Webber, and James P. Reilly. 157 nm photodissociation of a complete set of dipeptide ions containing C-terminal arginine. *J. Am. Soc. Mass. Spectrom.*, 24(5):675–683, 2013.

[153] Nathaniel Webber, Yi He, and James P. Reilly. 157 Nm Photodissociation of Dipeptide Ions Containing N-Terminal Arginine. *Journal of the American Society for Mass Spectrometry*, 25(2):196–203, 2014.

[154] Jared B Shaw, Wenzong Li, Dustin D Holden, Yan Zhang, Jens Griep-Raming, Ryan T Fellers, Bryan P Early, Paul M Thomas, Neil L Kelleher, and Jennifer S Brodbelt. Complete protein characterization using top-down mass spectrometry and ultraviolet photodissociation. *J. Am. Chem. Soc.*, 135(34):12646–51, aug 2013.

[155] Jared B Shaw, Errol W Robinson, and Ljiljana Paša-Tolić. Vacuum Ultraviolet Photodissociation and Fourier Transform-Ion Cyclotron Resonance (FT-ICR) Mass Spectrometry: Revisited. *Anal. Chem.*, 88(6):3019–3023, 2016.

[156] Ahmed Saeed, Federico Floris, Ulla Andersson, Irina Pikuleva, A. Lovgren-Sandblom, Maria Bjerke, Martin Paucar, Anders Wallin, Per Svenningsson, and I. Bjorkhem. 7$\alpha$-hydroxy-3-oxo-4-cholestenoic acid in cerebrospinal fluid reflects the integrity of the blood-brain barrier. *J. Lipid Res.*, 55(2):313–318, feb 2014.

[157] Christopher P Rüger, Theo Schwemer, Martin Sklorz, Peter B O'Connor, Mark P Barrow, and Ralf Zimmermann. Comprehensive chemical comparison of fuel composition and aerosol particles emitted from a ship diesel engine by gas chromatography atmospheric pressure chemical ionisation ultra-high

resolution mass spectrometry with improved data processing routines. *Eur. J. Mass Spectrom.*, 23(1):28–39, feb 2017.

[158] Xiao Zheng, An Kang, Chen Dai, Yan Liang, Tong Xie, Lin Xie, Yin Peng, Guangji Wang, and Haiping Hao. Quantitative analysis of neurochemical panel in rat brain and plasma by liquid chromatography-tandem mass spectrometry. *Anal. Chem.*, 84(22):10044–10051, 2012.

[159] Till Gruendling, Steffen Weidner, Jana Falkenhagen, and Christopher Barner-Kowollik. Mass spectrometry in polymer chemistry: a state-of-the-art up-date. *Polym. Chem.*, 1(5):599, 2010.

[160] L Pasa-Tolic, C Masselon, R C Barry, Y Shen, and R D Smith. Proteomic analyses using an accurate mass and time tag strategy. *Biotechniques*, 37(4):621–636., 2004.

[161] Y. Shen, R. Zhao, M. E. Belov, T. P. Conrads, G. A. Anderson, K. Tang, L. Paša-Tolić, T. D. Veenstra, M. S. Lipton, H. R. Udseth, and R. D. Smith. Packed capillary reversed-phase liquid chromatography with high-performance electrospray ionization fourier transform ion cyclotron resonance mass spectrometry for proteomics. *Analytical Chemistry*, 73(8):1766–1775, 2001.

[162] Jeremiah D. Tipton, John C. Tran, Adam D. Catherman, Dorothy R. Ahlf, Kenneth R. Durbin, and Neil L. Kelleher. Analysis of intact protein isoforms by mass spectrometry. *J. Biol. Chem.*, 286(29):25451–25458, 2011.

[163] Fred W. McLafferty. Tandem mass spectrometric analysis of complex biological mixtures. *Int. J. Mass Spectrom.*, 212(1-3):81–87, 2001.

[164] Jeremiah D. Tipton, John C. Tran, Adam D. Catherman, Dorothy R. Ahlf, Kenneth R. Durbin, Ji Eun Lee, John F. Kellie, Neil L. Kelleher, Christopher L. Hendrickson, and Alan G. Marshall. Nano-LC FTICR tandem mass spectrometry for top-down proteomics: Routine baseline unit mass resolution of whole cell lysate proteins up to 72 kDa. *Anal. Chem.*, 84(5):2111–2117, 2012.

[165] Federico Floris, Maria van Agthoven, Lionel Chiron, Andrew J. Soulby, Christopher A. Wootton, Yuko P. Y. Lam, Mark P. Barrow, Marc-André Delsuc, and Peter B. O'Connor. 2D FT-ICR MS of Calmodulin: A Top-Down and Bottom-Up Approach. *J. Am. Soc. Mass Spectrom.*, 27(9):1531–1538, sep 2016.

[166] Federico Floris, Claudio Vallotto, Lionel Chiron, Alice M. Lynch, Mark P. Barrow, Marc-André Delsuc, and Peter B. O'Connor. Polymer analysis in the second dimension: Preliminary studies for the characterization of polymers with 2d MS. *Anal. Chem.*, 89(18):9892–9899, aug 2017.

[167] Nadezda P. Sargaeva, Cheng Lin, and Peter B. O'Connor. Identification of aspartic and isoaspartic acid residues in amyloid $\beta$ peptides, including A$\beta$ 1-42, using electron-ion reactions. *Anal. Chem.*, 81(23):9778–9786, 2009.

[168] Pilar Perez Hurtado and Peter B O'Connor. Deamidation of collagen. *Anal. Chem.*, 84(6):3017–3025, 2012.

[169] Andrew J. Soulby, Jack W. Heal, Mark P. Barrow, Rudolf A. Roemer, and Peter B. O'Connor. Does deamidation cause protein unfolding? A top-down tandem mass spectrometry study. *Protein Sci.*, 24(5):850–860, 2015.

[170] Huilin Li, Stephen A. Wells, J. Emilio Jimenez-Roldan, Rudolf A. Römer, Yao Zhao, Peter J. Sadler, and Peter B. O'Connor. Protein flexibility is key to cisplatin crosslinking in calmodulin. *Protein Sci.*, 21(9):1269–1279, 2012.

[171] A Kumar, R R Ernst, and K Wüthrich. A two-dimensional nuclear Overhauser enhancement (2D NOE) experiment for the elucidation of complete proton-proton cross-relaxation networks in biological macromolecules. *Biochem. Biophys. Res. Commun.*, 95(1):1–6, 1980.

[172] Peter Pfandler, Geoffrey Bodenhausen, Jacques Rapin, Raymond Houriet, and Tino Gaumann. Two-dimensional Fourier transform ion cyclotron resonance mass spectrometry. *Chem. Phys. Lett.*, 138(2):195–200, 1987.

[173] R N Bracewell. *The Fourier Transform and its Applications*. New York, 2000.

[174] Alan G. Marshall and Francis R Verdun. *Fourier Transforms in NMR, Optical, and Mass Spectrometry A User's Handbook*. Elsevier B.V., Amsterdam, 1990.

[175] Charles W Ross, William J Simonsick, and David J Aaserud. Application of stored waveform ion modulation 2D-FTICR MS/MS to the analysis of complex mixtures. *Anal. Chem.*, 74(18):4625–33, sep 2002.

[176] M.A. van Agthoven, Christopher A Wootton, Lionel Chiron, Marie-aude Coutouly, Andrew Soulby, Juan Wei, Mark P Barrow, Marc-andré Delsuc, Christian Rolando, and Peter B. O'Connor. Two-Dimensional Mass Spectrometry for Proteomics , a Comparative Study with Cytochrome c. *Anal. Chem.*, 88(8):4409–4417, 2016.

[177] Maria A. van Agthoven, Lionel Chiron, Marie-Aude Coutouly, Marc-André Delsuc, and Christian Rolando. Two-dimensional ECD FT-ICR mass spectrometry of peptides and glycopeptides. *Anal. Chem.*, 84(13):5589–95, jul 2012.

[178] Maria A. van Agthoven, Lionel Chiron, Marie-Aude Coutouly, Akansha Ashvani Sehgal, Philippe Pelupessy, Marc-André Delsuc, and Christian Rolando. Optimization of the discrete pulse sequence for two-dimensional FT-ICR mass spectrometry using infrared multiphoton dissociation. *Int. J. Mass Spectrom.*, 370:114–124, sep 2014.

[179] S. H. Reiger. Starlight Scintillation and Atmospheric Turbulence. *Astron. J.*, 68(6):395–406, 1963.

[180] Gareth a Morris. Systematic sources of signal irreproducibility and t1 noise in high-field NMR spectrometers. *J. Magn. Reson.*, 100:316–328, 1992.

[181] J.A. Cadzow and M.-M. Wu. Analysis of transient data in noise. *IEE Proc. Part. F*, 134(1):69, 1987.

[182] Maria a van Agthoven, Marie-Aude Coutouly, Christian Rolando, and Marc-André Delsuc. Two-dimensional Fourier transform ion cyclotron resonance mass spectrometry: reduction of scintillation noise using Cadzow data processing. *Rapid comm. mass. spec.*, 25(11):1609–1616, jun 2011.

[183] Lionel Chiron, Maria A. van Agthoven, Bruno Kieffer, Christian Rolando, and Marc-André Delsuc. Efficient denoising algorithms for large experimental datasets and their applications in Fourier transform ion cyclotron resonance mass spectrometry. *Proc. Natl. Acad. Sci. U.S.A.*, 111(4):1385–90, jan 2014.

[184] Maria A. van Agthoven, Marc-André Delsuc, Geoffrey Bodenhausen, and Christian Rolando. Towards analytically useful two-dimensional Fourier transform ion cyclotron resonance mass spectrometry. *Anal. Bioanal.Chem.*, 405(1):51–61, oct 2012.

[185] Charles W. Ross, Shenheng Guan, Peter B. Grosshans, Tom L. Ricca, and Alan G. Marshall. Two-dimensional Fourier transform ion cyclotron resonance mass spectrometry/mass spectrometry with stored-waveform ion radius modulation. *J. Am. Chem. Soc.*, 115(17):7854–7861, aug 1993.

[186] Guillaume van der Rest and Alan G. Marshall. Noise analysis for 2D tandem Fourier transform ion cyclotron resonance mass spectrometry. *Int. J. Mass Spectrom.*, 210-211:101–111, sep 2001.

[187] Maria A. van Agthoven and Peter B. O'Connor. Two-dimensional mass spectrometry in a linear ion trap, anin silicomodel. *Rapid Commun. Mass Spectrom.*, 31(8):674–684, mar 2017.

[188] Jae C Schwartz, Michael W Senko, and John E.P Syka. A two-dimensional quadrupole ion trap mass spectrometer. *J. Am. Soc. Mass. Spectrom.*, 13(6):659–669, 2002.

[189] Raymond E March. An Introduction to Quadrupole Ion Trap Mass Spectrometry An Introduction to Quadrupole Ion Trap Mass Spectrometry. *J. Mass Spectrom.*, 9888(April 1997), 2016.

[190] F. Floris, L. Chiron, M.-A. Delsuc, M. P. Barrow, and P. B. O'Connor. Top-down deep sequencing of ubiquitin using 2d ms. Submitted to Anal. Chem., 2018.

[191] Federico Floris, L. Chiron, M.-A. Delsuc, M. P. Barrow, and P. B. O'Connor. Application of tandem two-dimensional mass spectrometry for top-down deep sequencing of calmodulin. Submitted to J. Am. Soc. Mass Spectrom., 2017.

[192] Lionel Chiron, Marie-Aude Coutouly, Jean-Philippe Starck, Christian Rolando, and Marc-André Delsuc. SPIKE a Processing Software dedicated to Fourier Spectroscopies. *arXiv:1608.06777 [physics.comp-ph]*, aug 2016.

[193] Dominique Tramesel, Vincent Catherinot, and Marc-André Delsuc. Modeling of NMR processing, toward efficient unattended processing of NMR experiments. *J. Magn. Reson.*, 188(1):56–67, sep 2007.

[194] Federico Floris, Maria van Aghtoven, Lionel Chiron, Christopher A. Wootton, P. Y. Lam, Mark P. Barrow, Marc-André Delsuc, and Peter B. O'Connor. Bottom-up two-dimensional electron-capture dissociation mass spectrometry of calmodulin. *J. Am. Soc. Mass Spectrom*, 2017.

[195] Maria A. van Agthoven, Marc-André Delsuc, and Christian Rolando. Two-dimensional FT-ICR/MS with IRMPD as fragmentation mode. *Int. J. Mass Spectrom.*, 306(2-3):196–203, sep 2011.

[196] Maria A. van Agthoven, Mark P. Barrow, Lionel Chiron, Marie-Aude Coutouly, David Kilgour, Christopher a Wootton, Juan Wei, Andrew Soulby, Marc-André Delsuc, Christian Rolando, and Peter B. O'Connor. Differentiating Fragmentation Pathways of Cholesterol by Two-Dimensional Fourier Transform Ion Cyclotron Resonance Mass Spectrometry. *J. Am. Soc. Mass Spectrom.*, 26(12):2105–2114, dec 2015.

[197] Ester Marotta, Roberta Seraglia, Fabio Fabris, and Pietro Traldi. Atmospheric pressure photoionization mechanisms. *Int. J. Mass Spectrom.*, 228(2):841 – 849, aug 2003. Special Issue: In honour of Helmut Schwarz.

[198] Alexander Rich and F.H.C. Crick. The molecular structure of collagen. *J. Mol. Biol.*, 3(5):483–IN4, 1961.

[199] J Bella, M Eaton, B Brodsky, and H M Berman. Crystal and molecular structure of a collagen-like peptide at 1.9 A resolution. *Science (New York, N.Y.)*, 266(5182):75–81, 1994.

[200] Hayley Simon, Maria A. van Agthoven, Pui Yiu Lam, Federico Floris, Lionel Chiron, Marc-Andre Delsuc, Christian Rolando, Mark Barrow, and Peter O' Connor. Uncoiling Collagen: A Multidimensional Mass Spectrometry Study. *Analyst*, 2015.

[201] AR DJR Means. Calmodulin - an intracellular calcium receptor. *Nature*, 285:73–77, 1980.

[202] Frits C Stevens. Calmodulin: an introduction. *Can. J. Biochem. Cell Biol.*, 61(5):906–910, 1982.

[203] M Zhang, T Tanaka, and M Ikura. Calcium-induced conformational transition revealed by the solution structure of apo calmodulin. *Nat. Struct. Biol.*, 2(9):758–767, 1995.

[204] Saša M. Miladinović, Cynthia J. Kaeser, Matthias M. Knust, and Charles L. Wilkins. Tandem Fourier transform mass spectrometry of block and random copolymers. *Int. J. Mass Spectrom.*, 301(1-3):184–194, 2011.

[205] Juan Wei, Anthony Bristow, Eileen McBride, David Kilgour, and Peter B. O'Connor. d--tocopheryl polyethylene glycol 1000 succinate: A view from FTICR MS and tandem MS. *Anal. Chem.*, 86(3):1567–1574, feb 2014.

[206] Pilar Perez Hurtado, Pui Yiu Lam, David Kilgour, Anthony Bristow, Eileen McBride, and Peter B. O'Connor. Use of high resolution mass spectrometry for analysis of polymeric excipients in drug delivery formulations. *Anal. Chem.*, 84(20):8579–8586, 2012.

[207] H Kato, Y Nagai, K Yamamoto, and Y Sakabe. Determination of polysorbates in foods by colorimetry with confirmation by infrared spectrophotometry, thin-layer chromatography, and gas chromatography. *J. Assoc. Off. Anal. Chem.*, 72(1):27–29, 1989.

[208] Fabrice Bray, Julien Bouclon, Lionel Chiron, Matthias Witt, Marc-André Delsuc, and Christian Rolando. Nonuniform sampling acquisition of two-dimensional fourier transform ion cyclotron resonance mass spectrometry for increased mass resolution of tandem mass spectrometry precursor ions. *Analytical Chemistry*, 89(17):8589–8593, aug 2017.

[209] Huilin Li, Tzu Yung Lin, Steve L. Van Orden, Yao Zhao, Mark P. Barrow, Ana M. Pizarro, Yulin Qi, Peter J. Sadler, and Peter B. O'Connor. Use of top-down and bottom-up fourier transform ion cyclotron resonance mass spectrometry for mapping calmodulin sites modified by platinum anticancer drugs. *Anal. Chem.*, 83:9507–9515, 2011.

[210] R.A. Zubarev, K.F. Haselmann, B. Budnik, F. Kjeldsen, and F. Jensen. Towards an understanding of the mechanism of electron-capture dissociation: a historical perspective and modern ideas. *European Journal of Mass Spectrometry*, 8(5):337–349, 2002.

[211] Lloyd M Smith and Neil L Kelleher. Proteoform: a single term describing protein complexity. *Nat. Methods*, 10(3):186–7, 2013.

[212] Jürgen Cox and Matthias Mann. Is Proteomics the New Genomics? *Cell*, 130(3):395–398, 2007.

[213] P H O'Farrell. High Resolution Two-Dimensional Electrophoresis of Proteins. *J. Biol. Chem.*, 250(10):4007–4021, 1975.

[214] Neil L. Kelleher. Top-Down Proteomics. *Analytical Chemistry*, 76(11):196 A–203 A, 2004.

[215] Neil L. Kelleher, Michael W. Senko, Siegel M. Marshall, and Fred W. McLafferty. Unit Resolution Mass Spectra of 112 kDa Molecules with 3 Da Accuracy. *J. Am. Soc. Mass Spectrom.*, 8:380–383, 1997.

[216] Peter Pfandler and Tino Gaumann. Broad-Band Two-Dimensional Fourier Transform Ion Cyclotron Resonance. *J. Am. Chem. Soc*, 110(17):5625–5628, 1988.

[217] Shenheng Guan and Patrick R. Jones. A theory for two-dimensional Fourier-transform ion cyclotron resonance mass spectrometry. *J. Chem. Phys*, 91(9):5291–5295, 1989.

[218] Mike Bensimon. A method to generate phase continuity in two-dimensional Fourier Transform Ion Cyclotron Resonance Mass Spectrometry. *Chem. Phys. Lett.*, 157(1):97–100, 1989.

[219] Youri O Tsybin, Matthias Witt, Gökhan Baykut, Frank Kjeldsen, and Per Håkansson. Combined infrared multiphoton dissociation and electron capture dissociation with a hollow electron beam in Fourier transform ion cyclotron resonance mass spectrometry. *Rapid comm. mass spectrom.*, 17(15):1759–68, jan 2003.

[220] Owen S. Skinner, Fred W. McLafferty, and Kathrin Breuker. How ubiquitin unfolds after transfer into the gas phase. *J. Am. Soc. Mass. Spectrom.*, 23(6):1011–1014, 2012.

[221] Kenneth R Durbin, Owen S Skinner, Ryan T Fellers, and Neil L Kelleher. Analyzing internal fragmentation of electrosprayed ubiquitin ions during beam-type collisional dissociation. *J. Am. Soc. Mass. Spectrom.*, 26(5):782–787, 2015.

[222] Yingying Huang, Joseph M. Triscari, George C. Tseng, Ljiljana Pasa-Tolic, Mary S. Lipton, Richard D. Smith, and Vicki H. Wysocki. Statistical characterization of the charge state and residue dependence of low-energy CID peptide dissociation patterns. *Anal. Chem.*, 77(18):5800–5813, 2005.

[223] Michael J. Bogan and George R. Agnes. Poly(ethylene glycol) doubly and singly cationized by different alkali metal ions: Relative cation affinities and cation-dependent resolution in a quadrupole ion trap mass spectrometer. *Journal of the American Society for Mass Spectrometry*, 13(2):177–186, 2002.

[224] Na Cao and Si Shen Feng. Doxorubicin conjugated to d-??-tocopheryl polyethylene glycol 1000 succinate (TPGS): Conjugation chemistry, characterization, in vitro and in vivo evaluation. *Biomaterials*, 29(28):3856–3865, 2008.

[225] Yuanyuan Guo, Jun Luo, Songwei Tan, Ben Oketch Otieno, and Zhiping

Zhang. The applications of Vitamin e TPGS in drug delivery. *Eur. J. Pharm. Sci.*, 49(2):175–186, 2013.

[226] Gene Hart-Smith and Christopher Barner-Kowollik. Contemporary mass spectrometry and the analysis of synthetic polymers: Trends, techniques and untapped potential. *Macromol. Chem. Phys.*, 211(14):1507–1529, 2010.

[227] Rebecca H. Wills, Manuela Tosin, and Peter B. O'Connor. Structural characterization of polyketides using high mass accuracy tandem mass spectrometry. *Anal. Chem.*, 84(20):8863–8870, 2012.

[228] Peter B. O'Connor and Fred W. McLafferty. Oligomer Characterization of 4-23 kDa Polymers by Electrospray Fourier Transform Mass Spectrometry. *J. Am. Chem. Soc.*, 117(51):12826–12831, dec 1995.

[229] Junkan Song, Jan W Van Velde, Luc L T Vertommen, Donald F Smith, Ron M A Heeren, and Oscar F Van Den Brink. End-Group Analysis of Methacrylic ( Co ) polymers by LC-ESI-MS 2. *Macromolecules*, 44:1319–1326, 2011.

[230] Malgorzata A. Kaczorowska and Helen J. Cooper. Characterization of Polyphosphoesters by Fourier Transform Ion Cyclotron Resonance Mass Spectrometry. *J. Am. Soc. Mass. Spectrom.*, 20(12):2238–2247, 2009.

[231] Francesca Bennet, Gene Hart-Smith, Till Gruendling, Thomas P. Davis, Philip J. Barker, and Christopher Barner-Kowollik. Degradation of poly(methyl methacrylate) model compounds under extreme environmental conditions. *Macromol. Chem. Phys.*, 211(10):1083–1097, may 2010.

[232] L. Mu and P. H. Seow. Application of TPGS in polymeric nanoparticulate drug delivery system. *Colloids Surf., B*, 47(1):90–97, 2006.

[233] L. Y. Kong, B. G. Su, Z. B. Bao, H. B. Xing, Y. W. Yang, and Q. L. Ren. Direct quantification of mono- and di-d-$\alpha$-tocopherol polyethylene glycol 1000 succinate by high performance liquid chromatography. *J. Chromatogr. A*, 1218(48):8664–8671, 2011.

[234] Jingyuan Wang, Baogen Su, Huabin Xing, Yiwen Yang, and Qilong Ren. Preparation and Characterization of Mono- and Di-d-a-Tocopheryl Polyethylene Glycol 1000 Succinate Jingyuan. *Journal of Polymer Science*, 119:3026–3033, 2011.

[235] Anne Christiansen, Thomas Backensfeld, Silke Kuhn, and Werner Weitschies. Investigating the Stability of the Nonionic Surfactants Tocopheryl Polyethylene Glycol Succinate and Sucrose Laurate by HPLCâMS, DAD, and CAD. *J. Pharm. Sci.*, 100:1773–1782, 2011.

[236] Deborah L. McKean, Amadeo J. Pesce, and Winston Koo. Analysis of polysorbate and its polyoxyethylated metabolite. *Anal. Biochem.*, 161(2):348–351, 1987.

[237] Folahan O Ayorinde, Slair V Gelain, James H Johnson, and Lily W Wan. Analysis of some commercial polysorbate formulations using matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. *Rapid Commun. Mass Spectrom.*, 14(22):2116–2124, 2000.

[238] Vikram S. Nayak, Zhijun Tan, Peter M. Ihnat, Reb J. Russell, and Michael J. Grace. Evaporative light scattering detection based HPLC method for the determination of polysorbate 80 in therapeutic protein formulations. *J. Chromatogr. Sci.*, 50(1):21–25, 2012.

[239] R. Zhang, Y. Wang, L. Tan, H. Y. Zhang, and M. Yang. Analysis of polysorbate 80 and its related compounds by RP-HPLC with ELSD and MS detection. *J. Chromatogr. Sci.*, 50(7):598–607, 2012.

[240] Yejia Li, Jessica N Hoskins, Subramanya G Sreerama, Michael A Grayson, and Scott M Grayson. The identification of synthetic homopolymer end groups and verification of their transformations using MALDI-TOFmass spectrometry. *J. Mass Spectrom.*, 45(6):587–611, 2010.

[241] G J Van Rooij, M C Duursma, R M A Heeren, J J Boon, and C G De Koster. G. J. van Rooij, C. G. de Koster M. C. Duursma, R. M. A. Heeren,. *J. Am. Soc. Mass Spectrom*, 7:449–457, 1996.

[242] S Koster, M C Duursma, J J Boon, M W F Nielen, C G de Koster, and R M A Heeren. Structural analysis of synthetic homo- and copolyesters by electrospray ionization on a Fourier transform ion cyclotron resonance mass spectrometer. *J. Mass Spectrom.*, 35(6):739–748, 2000.

[243] Sander Koster, MC Duursma, JJ Boon, and RM Heeren. Endgroup determination of synthetic polymers by electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry. *J. Am. Soc. Mass. Spectrom.*, 11(6):536–43, 2000.

[244] R.P. Lattimer. Tandem mass spectrometry of poly(ethylene glycol) proton- and deuteron-attachment ions. *Int. J. Mass Spectrom. Ion Processes*, 116(1):23–36, jul 1992.

[245] Blas Cerda, David Horn, Kathrin Breuker, Barry Carpenter, and Fred McLafferty. Electron capture dissociation of multiply-charged oxygenated cations. A nonergodic process. *Eur. J. Mass Spectrom.*, 5:335, 1999.

[246] Blas A. Cerda, Kathrin Breuker, David M. Horn, and Fred W. McLafferty. Charge/radical site initiation versus coulombic repulsion for cleavage of multiply charged ions. Charge solvation in poly(alkene glycol) ions. *J. Am. Soc. Mass. Spectrom.*, 12(5):565–570, 2001.

[247] Agnieszka Sawicka, Piotr Skurski, Robert R. Hudgins, and Jack Simons. Model calculations relevant to disulfide bond cleavage via electron capture influenced

by positively charged groups. *J. Phys. Chem. B*, 107(48):13505–13511, dec 2003.

[248] Yury O. Tsybin, John P. Quinn, Oleg Yu Tsybin, Christopher L. Hendrickson, and Alan G. Marshall. Electron Capture Dissociation Implementation Progress in Fourier Transform Ion Cyclotron Resonance Mass Spectrometry. *J. Am. Soc. Mass. Spectrom.*, 19(6):762–771, 2008.

[249] A Varki, R Cummings, J Esko, H Freeze, G Hart, and J Marth. Essentials of glycobiology cold spring harbor laboratory press. *New York*, pages 66–84, 1999.

[250] Kazuaki Ohtsubo and Jamey D. Marth. Glycosylation in Cellular Mechanisms of Health and Disease. *Cell*, 126(5):855–867, 2006.

[251] Rune Matthiesen. Mass Spectrometry Data Analysis in Proteomics. *Methods in Molecular Biology*, 42(4):2007, 2007.

[252] Rolf Apweiler, Henning Hermjakob, and Nathan Sharon. On the frequency of protein glycosylation, as deduced from analysis of the SWISS-PROT database. *Biochim. Biophys. Acta, Gen. Subj.*, 1473(1):4–8, 1999.

[253] Robert G Spiro. Protein glycosylation: nature, distribution, enzymatic formation, and disease implications of glycopeptide bonds. *Glycobiology*, 12(4):43R–56R, 2002.

[254] J.A. Harper, J.S. Yuan, J.B. Tan, I. Visan, and C.J. Guidos. Notch signaling in development and disease. *Clin. Genet.*, 1(3):461–472, 2003.

[255] Ghislain Opdenakker, Chris Dillen, Pierre Fiten, Erik Martens, Ilse Van Aelst, Philippe E. Van den Steen, Inge Nelissen, Sofie Starckx, Francis J. Descamps, Jialiang Hu, Helene Piccard, Jo Van Damme, Mark R. Wormald, Pauline M. Rudd, and Raymond A. Dwek. Remnant epitopes, autoimmunity and glycosylation. *Biochim. Biophys. Acta, Gen. Subj.*, 1760(4):610–615, 2006.

[256] Markus Sperandio. Selectins and glycosyltransferases in leukocyte rolling in vivo. *FEBS Journal*, 273(19):4377–4389, 2006.

[257] Hudson H Freeze. Genetic defects in the human glycome. *Nat Rev Genet*, 7(7):537–551, jul 2006.

[258] Joseph Zaia. Mass Spectrometry and the Emerging Field of Glycomics. *Chem. Biol.*, 15(9):881–892, 2008.

[259] Harry Schachter. The joys of HexNAc. The synthesis and function of N- and O-glycan branches. *Glycoconjugate Journal*, 17(7-9):465–483, 2000.

[260] Suzan Wopereis, Dirk J. Lefeber, Éva Morava, and Ron A. Wevers. Mechanisms in protein O-glycan biosynthesis and clinical and molecular aspects of protein O-glycan biosynthesis defects: A review. *Clin. Chem.*, 52(4):574–600, 2006.

[261] A.L. Tarentino and T.H. Plummer Jr. Deglycosylation of asparagine-linked glycans by peptide:N-glycosidase F. *Trends Glycosci. Glycotechnol.*, 24(17):4665–71, 1985.

[262] Anthony L Tarentino and Thomas H Plummer. Enzymatic deglycosylation of asparagine-linked glycans: Purification, properties, and specificity of oligosaccharide-cleaving enzymes from Flavobacterium meningosepticum. *Methods in enzymology*, 230:44–57, 1994.

[263] Seiichi Takasaki, Tsuguo Mizuochi, and Akira Kobata. [17] hydrazinolysis of asparagine-linked sugar chains to produce free oligosaccharides. In *Complex Carbohydrates Part D*, pages 263–268. Elsevier, 1982.

[264] D R Wing, T W Rademacher, M C Field, R a Dwek, B Schmitz, G Thor, and M Schachner. Use of large-scale hydrazinolysis in the preparation of N-linked oligosaccharide libraries: application to brain tissue. *Glycoconjugate J.*, 9(6):293–301, 1992.

[265] T Patel, J Bruce, a Merry, C Bigge, M Wormald, a Jaques, and R Parekh. Use of hydrazine to release in intact and unreduced form both N- and O-linked oligosaccharides from glycoproteins. *Biochemistry*, 32(2):679–693, 1993.

[266] David J. Harvey. Derivatization of carbohydrates for analysis by chromatography; electrophoresis and mass spectrometry. *J. Chromatogr. B Analyt. Technol. Biomed. Life Sci.*, 879(17-18):1196–1225, 2011.

[267] S. Hunter Walker, Laura M. Lilley, Monica F. Enamorado, Daniel L. Comins, and David C. Muddiman. Hydrophobic derivatization of N-linked glycans for increased ion abundance in electrospray ionization mass spectrometry. *J. Am. Soc. Mass. Spectrom.*, 22(8):1309–1317, 2011.

[268] Ionel Ciucanu and Catherine E. Costello. Elimination of Oxidative Degradation during the per-O-Methylation of Carbohydrates. *J. Am. Chem. Soc.*, 125(52):16213–16219, 2003.

[269] Joseph Zaia. Mass spectrometry of oligosaccharides. *Mass Spectrom. Rev.*, 23(3):161–227, 2004.

[270] J. Lemoine, F. Chirat, and B. Domon. Structural analysis of derivatized oligosaccharides using post-source decay matrix-assisted laser desorption/ionization mass spectrometry. *Journal of Mass Spectrometry*, 31(8):908–912, 1996.

[271] David J. Harvey. Identification of protein-bound carbohydrates by mass spectrometry. *Proteomics*, 1(2):311–328, 2001.

[272] David J. Harvey. Structural determination of N-linked glycans by matrix-assisted laser desorption/ionization and electrospray ionization mass spectrometry. *Proteomics*, 5(7):1774–1786, 2005.

[273] Youmie Park and Carlito B Lebrilla. Application of Fourier transform ion cyclotron resonance mass spectrometry to oligosaccharides. *Mass Spectrom Rev*, 24(July 2003):232–264, 2005.

[274] Joseph Zaia. Mass Spectrometry and Glycomics. *OMICS*, 14(4):401–418, aug 2010.

[275] Wen Zhou and Kristina Hakansson. Structural characterization of carbohydrates by fourier transform tandem mass spectrometry. *Curr. Proteomics*, 8(4):297–308, dec 2011.

[276] Manfred Wuhrer. Glycomics using mass spectrometry. *Glycoconjugate J.*, 30(1):11–22, 2013.

[277] D. J. Harvey. Quantitative aspects of the matrix-assisted laser desorption mass spectrometry of complex oligosaccharides. *Rapid Commun. Mass Spectrom.*, 7(7):614–619, jul 1993.

[278] Cheng Zhao, Bo Xie, Shiu Yung Chan, Catherine E. Costello, and Peter B. O'Connor. Collisionally Activated Dissociation and Electron Capture Dissociation Provide Complementary Structural Information for Branched Permethylated Oligosaccharides. *J. Am. Soc. Mass. Spectrom.*, 19(1):138–150, 2008.

[279] Julie T Adamson and Kristina Håkansson. Electron Capture Dissociation of Oligosaccharides Ionized with Alkali, Alkaline Earth, and Transition Metals. *Anal. Chem.*, 79(7):2901–2910, 2007.

[280] Jeremy J. Wolff, I. Jonathan Amster, Lianli Chi, and Robert J. Linhardt. Electron Detachment Dissociation of Glycosaminoglycan Tetrasaccharides. *J. Am. Soc. Mass. Spectrom.*, 18(2):234–244, 2007.

[281] James R. Kornacki, Julie T. Adamson, and Kristina Håkansson. Electron Detachment Dissociation of Underivatized Chloride-Adducted Oligosaccharides. *J. Am. Soc. Mass. Spectrom.*, 23(11):2031–2042, nov 2012.

[282] Xiang Yu, Yiqun Huang, Cheng Lin, and Catherine E. Costello. Energy-dependent electron activated dissociation of metal-adducted permethylated oligosaccharides. *Anal. Chem.*, 84(17):7487–7494, 2012.

[283] J. A. Hofmeister, G. E.; Zhou, Z.; Leary, Gretchen E. Hofmeister, Zhongrui Zhou, and Julie a. Leary. Linkage position determination in lithium-cationized disaccharides: tandem mass spectrometry and semiempirical calculations. *J. Am. Chem. Soc.*, 113(16):5964–5970, 1991.

[284] Mark T. Cancilla, Sharron G. Penn, James A. Carroll, and Carlito B. Lebrilla. Coordination of alkali metals to oligosaccharides dictates fragmentation behavior in matrix assisted laser desorption ionization/Fourier transform mass spectrometry. *J. Am. Chem. Soc.*, 118(28):6736–6745, jan 1996.

[285] E.W. Collings and S.C. Hart. Low temperature magnetic susceptibility and magnetization studies of some commercial austenitic stainless steels. *Cryogenics*, 19(9):521–530, sep 1979.

[286] Andrew Sunderland, Li Ju, David G Blair, Wayne McRae, and Howard Golden. Low magnetic susceptibility materials and applications in magnetic gradiometry. *Smart Mater. Struct.*, 18(9):095038, aug 2009.

# About the author

**Federico Floris** obtained his MSc at Alma mater studiorum, University of Bologna, after which he started working for Unilever R&D, Port Sunlight, until he started his Ph.D. at the University of Warwick in 2014. He is completing his Ph.D. under the supervision of Prof. Peter B. O'Connor, with a project focused on the development and applications of two-dimensional Fourier transform ion cyclotron resonance mass spectrumetry. During his years at Warwick, he also completed the Certificate for transferable skills in science, and became chair of the LGBT+ staff and postgraduate network. In 2016 he was featured as an LGBT role model for the University of Warwick.

## Publications

### During the Ph.D.

**Floris, F.** and O'Connor, P.B. Fundamentals of 2D FT-MS. In Fundamentals and Applications of Fourier Transform Mass Spectrometry; Schmitt-Kopplin, P. and Kanawati, B, Ed.; Elsevier, **2018**. *Chapter submitted.*

**Floris, F.**; Chiron L.; Delsuc M.-A.; Barrow P.M.; O'Connor, P.B. (2017): Application of tandem two-dimensional mass spectrometry for Top-Down deep sequencing of Calmodulin. *J. Am. Soc.*, **2018**, *Mass Spectrom. Article submitted.*

**Floris, F.**; Chiron L.; Delsuc M.-A.; Barrow P.M.; O'Connor, P.B.: Top-Down Deep Sequencing of Ubiquitin using 2D MS. *Anal. Chem.*, **2018**, *Article submitted.*

**Floris, F.**; van Aghtoven, M.A.; Chiron L.; Wootton A.C.; Lam P.Y.; Delsuc M.-A.; Barrow P.M.; O'Connor, P.B.: Bottom-up two-dimensional electron-capture dissociation mass spectrometry of calmodulin. *J. Am. Soc. Mass Spectrom.,* **2017**, 29 (1), 207-210.

**Floris, F.**; Vallotto, C.; Chiron L.; Lynch, A.; Delsuc M.-A.; Barrow P.M.; O'Connor, P.B. (2017): Polymer analysis in the second dimension: preliminary studies for the characterisation of polymers with 2D MS. *Anal. Chem.*, **2017**, 89 (18), 9892-9899.

**Floris, F.**; van Aghtoven, M.A.; Chiron L.; Wootton A.C.; Lam P.Y.; Delsuc M.-A.; Barrow P.M.; O'Connor, P.B. (2016): Two-dimensional Fourier Transform Ion Cyclotron Resonance Mass Spectrometry of Calmodulin: a top-down and bottom-up approach. *J. Am. Soc. Mass Spectrom.*, **2016**, 27(9), 1531–1538.

Simon, H., van Agthoven, M.A., Lam, P.Y., **Floris, F.**, Chiron, L., Delsuc, M.-A., Rolando, C., Barrow, M., O'Connor, P.: Uncoiling Collagen: a multidimensional mass spectrometry study. *Analyst.*, **2015**, 141, 157-165.

## During the M.Sc.

Saeed A.[1], **Floris F.**[1], Andersson U., Pikuleva I., Lövgren-Sandblom A., Bjerke M., Paucar M., Wallin A., Svenningsson P., Björkhem I., "7α-hydroxy-3-oxo-4-cholestenoic acid in cerebrospinal fluid reflects the integrity of the blood-brain barrier." *J Lipid Res.* **2014** Feb; 55(2):313-8. [1]Authors with equal contribution.

# A. Supporting Information for Chapter 3

# A.1. IRMPD 2D MS of Calmodulin

## A.1.1. Top-Down analysis of *c*almodulin

Table A.1.: Peak assignment for top-down 2D-IRMPD/FT-ICR MS Calmodulin.

| **Fragment ions** | **Charge** | **Theoreti*c*al** $m/z$ | **Observed** $m/z$ | **Error [ppm]** |
|---|---|---|---|---|
| $y_2$ | 1 | 218.149918 | 218.150115 | 0.90 |
| $b_2$ | 1 | 229.081899 | 229.082081 | 0.79 |
| $y_3$ | 1 | 319.197597 | 319.197628 | 0.10 |
| $b_3$ | 1 | 357.140477 | 357.140358 | -0.33 |
| PTEA-$H_2O$ | 1 | 381.176900 | 381.176642 | -0.58 |
| $y_4$-$H_2O$ | 1 | 432.227517 | 432.227401 | -0.27 |
| $y_4$ | 1 | 450.238082 | 450.237949 | -0.29 |
| $b_4$ | 1 | 470.224541 | 470.224312 | -0.49 |
| $b_5$-$H_2O$ | 1 | 553.261655 | 553.260604 | -1.90 |
| $y_5$-$H_2O$ | 1 | 563.268002 | 563.267836 | -0.29 |
| $b_5$ | 1 | 571.272220 | 571.271558 | -1.16 |
| $y_5$ | 1 | 581.278567 | 581.278309 | -0.44 |
| EFVQM-$H_2O$ | 1 | 617.275200 | 617.274524 | -1.09 |
| PTEAEL | 1 | 641.314100 | 641.313744 | -0.53 |
| $b_6$-$H_2O$ | 1 | 682.304249 | 682.302961 | -1.89 |
| $y_6$-NH3 | 1 | 692.310596 | 692.309784 | -1.17 |
| ? | 2 | | 696.847448 | |
| $y_6$ | 1 | 709.337145 | 709.337590 | 0.63 |
| QDMINE-NH3 | 1 | 714.276300 | 714.275576 | -1.04 |
| MIREAD | 1 | 716.339600 | 716.339328 | -0.36 |
| EEFVQM-$H_2O$ | 1 | 746.317800 | 746.317438 | -0.47 |
| NPTEAEL | 1 | 755.357000 | 755.357591 | 0.77 |
| ? | 2 | | 804.884127 | |
| $y_7$ | 1 | 808.405559 | 808.406023 | 0.57 |
| EMIREAD-$H_2O$ | 1 | 827.371600 | 827.368425 | -3.86 |
| $b_7$ | 1 | 829.357408 | 829.355787 | -1.95 |
| EMIREAD | 1 | 845.382200 | 845.380091 | -2.47 |
| PTEAELQD-$H_2O$ | 1 | 866.389000 | 866.388953 | -0.10 |
| b15 | 2 | 873.415060 | 873.415287 | 0.26 |
| $y_{15}$-$H_2O$ | 2 | 878.905293 | 878.904927 | -0.42 |
| ? | 2 | | 898.900820 | |
| ? | 1 | | 903.915715 | |
| ? | 1 | | 912.920464 | |

| | | | | |
|---|---|---|---|---|
| $y_8$ | 1 | 955.473973 | 955.473651 | -0.34 |
| $y_{17}$-$H_2O$ | 2 | 964.929497 | 964.930418 | 0.96 |
| ? | 2 | | 984.442158 | |
| ? | 6 | | 1,002.815433 | |
| $b_{37}$-$H_2O$ | 4 | 1,035.758406 | 1,035.757729 | -0.65 |
| ? | 1 | | 1,044.466336 | |
| $b_9$-$H_2O$ | 1 | 1,052.489485 | 1,052.488617 | -0.82 |
| $b_9$ | 1 | 1,070.500050 | 1,070.494920 | -4.79 |
| $y19$-$H_2O$ | 2 | 1,078.985001 | 1,078.984842 | -0.15 |
| $b_{39}$-$2H_2O$ | 4 | 1,081.284788 | 1,081.286864 | 1.92 |
| $b_{39}$-$H_2O$ | 4 | 1,085.787430 | 1,085.788762 | 1.23 |
| $b_{39}$ | 4 | 1,090.290071 | 1,090.289966 | -0.10 |
| $b_{40}$-$H_2O$ | 4 | 1,100.042796 | 1,100.043707 | 0.83 |
| $b_{50}$-NH3 | 5 | 1,105.531262 | 1,105.533070 | 1.64 |
| $b_{51}$-NH3 | 5 | 1,131.739359 | 1,131.743737 | 3.87 |
| $b_{41}$-$H_2O$ | 4 | 1,132.057440 | 1,132.056631 | -0.71 |
| $b_{41}$ | 4 | 1,136.560081 | 1,136.556569 | -3.09 |
| $b_{10}$ | 1 | 1,141.537164 | 1,141.534267 | -2.54 |
| $b_{52}$-NH3 | 5 | 1,154.356172 | 1,154.357187 | 0.88 |
| ? | 4 | | 1,156.564910 | |
| $b_{42}$-NH3 | 4 | 1,160.814176 | 1,160.822771 | 7.40 |
| $b_{42}$ | 4 | 1,165.070813 | 1,165.072374 | 1.34 |
| MH | | | 1,199.924136 | |
| $b_{45}$-$H_2O$ | 4 | 1,242.353931 | 1,242.357154 | 2.59 |
| $b_{45}$ | 4 | 1,246.856573 | 1,246.854922 | -1.32 |
| ? | 2 | | 1,248.638688 | |
| ? | 2 | | 1,257.139891 | |
| ? | 1 | | 1,277.617715 | |
| ? | 1 | | 1,299.609740 | |
| $b_{22}$ | 2 | 1,299.615946 | 1,299.617500 | 1.20 |
| ? | 1 | | 1,317.554812 | |
| ? | 2 | | 1,361.650768 | |
| ? | 2 | | 1,370.195414 | |
| ? | 2 | | 1,378.707026 | |
| $b_{24}$ | 2 | 1,385.640150 | 1,385.638665 | -1.07 |
| ? | 1 | | 1,611.810091 | |
| $y_{15}$ | 1 | 1,774.813874 | 1,774.813333 | -0.30 |
| | | Average Error | | -0.18 |
| | | Standard deviation | | 1.85 |

### A.1.2. Bottom-Up analysis of Calmodulin

Table A.3.: Bottom-up 2D-IRMPD/FT-ICR MS CaM, peak assignment for ion at 403.215463 *m/z*.

| **Ion:** $[31\_37]^{2+}$ | | | |
|---|---|---|---|
| **Fragment** | **Theoretical** *m/z* | **Observed** *m/z* | **Error [ppm]** |
| $y$ | 175.118952 | 175.118877 | -0.43 |
| $b_2$-$H_2O$ | 225.123369 | 225.123466 | 0.43 |
| $y_2$ | 306.159437 | 306.159510 | 0.24 |
| $b_4$-$2H_2O$ | 365.181946 | 365.182217 | 0.74 |
| $b_4$-$H_2O$ | 383.192511 | 383.192650 | 0.36 |
| MH-$H_2O$ | 394.210181 | 394.208981 | -3.04 |
| $[MH_2]^{2+}$ | 403.215463 | 403.215860 | 0.98 |
| $y_3$ | 405.227851 | 405.227721 | -0.32 |
| $b_5$-$H_2O$ | 482.260925 | 482.260791 | -0.28 |
| $y_4$ | 506.275530 | 506.275810 | 0.55 |
| $y_5$-$H_2O$ | 545.286428 | 545.286426 | 0.00 |
| $y_5$ | 563.296993 | 563.297137 | 0.26 |
| $b_6$-$2H_2O$ | 595.290845 | 595.290598 | -0.42 |
| $b_6$-$H_2O$ | 613.301410 | 613.301637 | 0.37 |
| | Average Error | | -0.04 |
| | Standard deviation | | 0.93 |

Table A.4.: Bottom-up 2D-IRMPD/FT-ICR MS CaM, peak assignment for ion at 547.235832 *m/z*.

| **Ion:** $[78\_86]^{2+}$ | | | |
|---|---|---|---|
| **Fragment** | **Theoretical** *m/z* | **Observed** *m/z* | **Error [ppm]** |
| $y$ | 175.118952 | 175.118983 | 0.18 |
| $b_2$ | 217.081898 | 217.081864 | -0.16 |
| $y_2$ | 288.203016 | 288.202904 | -0.39 |
| $b_3$-$H_2O$ | 314.098276 | 314.098260 | -0.05 |
| $b_3$ | 332.108841 | 332.108772 | -0.21 |
| $y_3$ | 417.245609 | 417.245397 | -0.51 |
| $[MH_2$-$H_2O]^{2+}$ | 538.230550 | 538.231097 | 1.02 |
| $y_4$ | 546.288202 | 546.287524 | -1.24 |
| $[MH_2]^{2+}$ | 547.235832 | 547.235367 | -0.85 |
| $b_5$ | 548.183462 | 548.183563 | 0.18 |
| $b_6$-$2H_2O$ | 641.204926 | 641.203951 | -1.52 |

| | | | |
|---|---|---|---|
| $b_6$-$H_2O$ | 659.215491 | 659.214806 | -1.04 |
| $y_5$ | 675.330795 | 675.329826 | -1.43 |
| $b_6$ | 677.226056 | 677.225820 | -0.35 |
| $y_6$ | 762.362824 | 762.363086 | 0.34 |
| $b_7$-$2H_2O$ | 770.247519 | 770.247237 | -0.37 |
| $b_7$-$H_2O$ | 788.258084 | 788.258059 | -0.03 |
| $b_7$ | 806.268649 | 806.267984 | -0.82 |
| $y_7$-$H_2O$ | 859.379202 | 859.377517 | -1.96 |
| $b_8$-$3H_2O$ | 865.321018 | 865.324568 | 4.10 |
| $y_7$ | 877.389767 | 877.388923 | -0.96 |
| $b_8$-$2H_2O$ | 883.331583 | 883.332027 | 0.50 |
| $b_8$-$H_2O$ | 901.342148 | 901.342227 | 0.09 |
| $b_8$ | 919.352713 | 919.352639 | -0.08 |
| | Average Error | | -0.23 |
| | Standard deviation | | 1.14 |

Table A.5.: Bottom-up 2D-IRMPD/FT-ICR MS CaM, peak assignment for ion at 585.628447 *m/z*.

| **Ion:** $[91\_106]^{3+}$ | | | |
|---|---|---|---|
| **Fragment** | **Theoretical** *m/z* | **Observed** *m/z* | **Error [ppm]** |
| ? | | 441.767966 | |
| ? | | 443.783189 | |
| $y_4$ | 488.282723 | 488.282700 | -0.05 |
| ? | | 519.799132 | |
| $[MH_3]^{3+}$ | 585.628447 | 585.628184 | -0.45 |
| $d[MH_3]^{3+}$ | 585.956452 | 585.955704 | -1.28 |
| $y_6$ | 646.351865 | 646.351900 | 0.05 |
| $y_{14}{}^{2+}$ | 755.362627 | 755.362600 | -0.04 |
| | Average Error | | -0.35 |
| | Standard deviation | | 0.49 |

Table A.6.: Bottom-up 2D-IRMPD/FT-ICR MS CaM, peak assignment for ion at 615.635268 *m/z*.

| **Ion:** $[14\_30]^{3+}$ | | | |
|---|---|---|---|
| **Fragment** | **Theoretical** *m/z* | **Observed** *m/z* | **Error [ppm]** |
| $y_3$ | 349.208161 | 349.208136 | -0.07 |
| $y_4$ | 462.292225 | 462.292275 | 0.11 |

| | | | |
|---|---|---|---|
| $b_5$-$2H_2O$ | 512.250360 | 512.250999 | 1.25 |
| $b_5$-$H_2O$ | 530.260925 | 530.261639 | 1.35 |
| $[y_{15}$-$2H_2O]^{3+}$ | 536.934989 | 536.936343 | 2.52 |
| $[y_{11}$-$2H_2O]^{2+}$ | 557.772385 | 557.773139 | 1.35 |
| $y_5$ | 563.339904 | 563.341754 | 3.28 |
| $[y_{11}$-$H_2O]^{2+}$ | 566.777668 | 566.777777 | 0.19 |
| $[MH_3$-$5H_2O]^{3+}$ | 585.617660 | 585.617477 | -0.31 |
| $[MH_3$-$4H_2O]^{3+}$ | 591.621181 | 591.621636 | 0.77 |
| $[MH_3$-$3H_2O]^{3+}$ | 597.624703 | 597.624851 | 0.25 |
| $y_6$-$H_2O$ | 602.350802 | 602.351181 | 0.63 |
| $[MH_3$-$2H_2O]^{3+}$ | 603.628225 | 603.629659 | 2.38 |
| $[MH_3$-$H_2O]^{3+}$ | 609.631746 | 609.631368 | -0.62 |
| $[MH]^{3+}$ | 615.635268 | 615.635618 | 0.57 |
| $y_6$ | 620.361367 | 620.361259 | -0.17 |
| $[y_{12}$-$3H_2O]^{2+}$ | 622.301310 | 622.300789 | -0.84 |
| $[y_{12}$-$2H_2O]^{2+}$ | 631.306592 | 631.306366 | -0.36 |
| $[y_{12}$-$H_2O]^{2+}$ | 640.311875 | 640.311903 | 0.04 |
| ${y_{12}}^{2+}$ | 649.317157 | 649.317692 | 0.82 |
| $[b_{13}$-$2H_2O]^{2+}$ | 674.296224 | 674.296382 | 0.23 |
| $[b_{13}$-$H_2O]^{2+}$ | 683.301507 | 683.299349 | -3.16 |
| $[y_{13}$-$2H_2O]^{2+}$ | 687.848624 | 687.847927 | -1.01 |
| $[y_{13}$-$H_2O]^{2+}$ | 696.853907 | 696.851851 | -2.95 |
| $[b_{14}$-$3H_2O]^{2+}$ | 721.832974 | 721.832555 | -0.58 |
| $[b_{14}$-$2H_2O]^{2+}$ | 730.838256 | 730.837826 | -0.59 |
| $[y_{14}$-$H_2O]^{2+}$ | 740.369921 | 740.370325 | 0.55 |
| | Average Error | | 0.21 |
| | Standard deviation | | 1.37 |

Table A.7.: Bottom-up 2D-IRMPD/FT-ICR MS CaM, peak assignment for ion at 782.380486 *m/z*.

| **Ion:** $[1\_13]^{2+}$ | | | |
|---|---|---|---|
| **Fragment** | **Theoretical** *m/z* | **Observed** *m/z* | **Error [ppm]** |
| [ppm] | | | |
| $y_2$ | 294.181218 | 294.181670 | 1.54 |
| $b_3$ | 357.140476 | 357.140078 | -1.11 |
| $y_3$-$H_2O$ | 405.213246 | 405.213674 | 1.06 |
| $y_3$ | 423.223811 | 423.223816 | 0.01 |
| $b_4$ | 470.224540 | 470.224526 | -0.03 |

| | | | |
|---|---|---|---|
| $y_4$ | 494.260925 | 494.260813 | -0.23 |
| $b_5$-$H_2O$ | 553.261653 | 553.261102 | -1.00 |
| $b_5$ | 571.272218 | 571.271978 | -0.42 |
| $y_5$ | 607.344989 | 607.346291 | 2.14 |
| TEEQIA-$H_2O$ | 654.309334 | 654.309487 | 0.23 |
| TEEQIA | 672.319899 | 672.320422 | 0.78 |
| $b_6$-$H_2O$ | 682.304246 | 682.304275 | 0.04 |
| ${y_{12}}^{2+}$ | 725.856646 | 725.855826 | -1.13 |
| $[MH_2\text{-}H_2O]^{2+}$ | 773.375203 | 773.375467 | 0.34 |
| $[MH_2]^{2+}$ | 782.380486 | 782.380127 | -0.46 |
| LTEEQIA | 785.403963 | 785.403437 | -0.67 |
| $b_7$-$H_2O$ | 811.346839 | 811.349151 | 2.85 |
| $b_7$ | 829.357404 | 829.357680 | 0.33 |
| $y_7$-$H_2O$ | 846.435595 | 846.435030 | -0.67 |
| $b_8$ | 957.415982 | 957.451126 | 36.71 |
| DQLTEEQIA-$H_2O$ | 1010.478920 | 1010.476830 | -2.07 |
| DQLTEEQIA | 1028.489485 | 1028.490346 | 0.84 |
| $b_9$-$2H_2O$ | 1034.478916 | 1034.478881 | -0.03 |
| $b_9$-$H_2O$ | 1052.489481 | 1052.488447 | -0.98 |
| $b_9$ | 1070.500046 | 1070.501445 | 1.31 |
| $y9$ | 1094.536431 | 1094.538171 | 1.59 |
| $b_{10}$-$H_2O$ | 1123.526595 | 1123.526826 | 0.21 |
| $b_{10}$ | 1141.537160 | 1141.534467 | -2.36 |
| DQLTEEQIAE | 1157.532079 | 1157.534665 | 2.23 |
| $b_{11}$ | 1270.579753 | 1270.580712 | 0.75 |
| $b_{12}$ | 1417.648167 | 1417.647512 | -0.46 |
| | Average Error | | 1.33 |
| | Standard deviation | | 6.56 |

### A.1.3. Parameters for the acquisition of the two-dimensional spectra

Table A.8.: Parameters used for 2D FT-ICR MS spectral acquisition.

| **Parameter** | **Experiment** | |
|---|---|---|
| | **TDP 2D FT-ICR MS** | **BUP FT-ICR MS** |
| **Acquisition mode** | | |
| Acquisition size [datapoints] | 4 | 512k |
| Horizontal low mass cut-off [$m/z$] | 147.41 | 147.41 |
| Horizontal high mass cut-off [$m/z$] | 3000 | 3000 |

| | | |
|---|---|---|
| Transient length, t2 [s] | 1.6777 | 0.2097 |
| Ion accumulation time [s] | 0.1 | 0.1 |
| Ion cooling time [s] | 0.1 | 0.25 |
| **Source (nanoESI)** | | |
| Capillary voltage [V] | 900 | 1300 |
| Capillary exit [V] | 200 | 200 |
| Deflector plate [V] | 220 | 220 |
| Funnel [V] | 150 | 150 |
| Skimmer [V] | 10 | 10 |
| Funnel RF [Vpp] | 200 | 200 |
| **Collision cell** | | |
| Collision voltage [V] | 0 | 0 |
| DC extract bias [V] | 0.6 | 0.6 |
| RF frequency [MHz] | 2 | 2 |
| Collision RF Amplitude [Vpp] | 1400 | 1400 |
| **Octopole** | | |
| Frequency [MHz] | 2 | 2 |
| RF Amplitude [Vpp] | 1400 | 1400 |
| **Transfer optics** | | |
| Time of Flight [s] | 0.0012 | 0.0012 |
| Frequency [MHz] | 4 | 4 |
| RF Amplitude [Vpp] | 350 | 350 |
| **ICR-cell** | | |
| Transfer exit Lens [V] | -15 | -15 |
| Analyser entrance [V] | -8 | -8 |
| Side kick [V] | 12 | 10 |
| Side kick offset [V] | -2.3 | -2.9 |
| Front Trap plate [V] | 0.51 | 0.69 |
| Back Trap plate [V] | 0.55 | 0.74 |
| Sweep Excitation power, P3 [%] | 16 | 16 |
| **2D-MS Pulse Program** | | |
| $t_1$ [s] | 1.2E-06 | 2.4E-06 |
| $t_1$ increment [s] | 1.2E-06 | 2.4E-06 |
| $P_1$ and $P_2$ pulse length [s] | 4.0E-07 | 4.0E-07 |
| $P_1$ and $P_2$ excitation power [Db] | 10.42 | 10.42 |
| Laser Power (25W) [%] | 70 | 70 |
| IR pulse length [s] | 0.3 | 0.3 |

Table A.2.: Peak assignment for bottom-up 2D-IRMPD/FT-ICR MS Calmodulin. For the precursor ions indicated with a * symbol, MS/MS analysis is available in tables A.3 to A.7 below.

| **Peptide** | **Theoretical** $m/z$ | **Observed** $m/z$ | **Error [ppm]** | **Peptide sequence** | **Charge** |
|---|---|---|---|---|---|
| 31_37* | 403.215464 | 403.215329 | -0.33 | ELGTVMR | 2 |
| 76_86 | 451.538131 | 451.537898 | -0.52 | MKDTDSEEEIR | 3 |
| 22_30 | 454.221999 | 454.222146 | 0.32 | DGDGTITTK | 2 |
| 14_21 | 478.739827 | 478.740266 | 0.92 | EAFSLFDK | 2 |
| 75_86 | 494.236452 | 494.236087 | -0.74 | KMKDTDSEEEIR | 2 |
| 87_90 | 522.267074 | 522.266960 | -0.22 | EAFR | 1 |
| 78_90 | 532.909393 | 532.910201 | 1.52 | DTDSEEEIREAFR | 3 |
| 78_86* | 547.235835 | 547.235994 | 0.29 | DTDSEEEIR | 2 |
| 91_106* | 585.628449 | 585.628964 | 0.88 | VFDKDGNGYISAAELR | 3 |
| d91_106 | 585.956454 | 585.956483 | 0.05 | VFDKDGNGYISAAELR | 3 |
| 14_30* | 615.635270 | 615.636243 | 1.58 | EAFSLFDKDGDGTITTK | 3 |
| 76_86 | 676.803559 | 676.804002 | 0.66 | MKDTDSEEEIR | 2 |
| 1_13* | 782.380488 | 782.380453 | -0.04 | Ac-ADQLTEEQIAEFK | 2 |
| 78_90 | 798.860452 | 798.856303 | -5.19 | DTDSEEEIREAFR | 2 |
| 106_125 | 801.062778 | 801.057993 | -5.97 | HVMTNLGEKLTDEEVDEMIR | 3 |
| 31_37 | 805.423652 | 805.422482 | -1.45 | ELGTVMR | 1 |
| d91_106 | 878.431044 | 878.429467 | -1.80 | VFDKDGNGYISAAELR | 2 |
| 14_30 | 922.949267 | 922.950409 | 1.24 | EAFSLFDKDGDGTITTK | 2 |
| 78_86 | 1093.464394 | 1093.461091 | -3.02 | DTDSEEEIR | 1 |
| 127_148 | 1245.543609 | 1245.534997 | -6.91 | EADIDGDGQVNYEEFVQMMTAK | 2 |
| Average Error | | | -0.94 | | |
| Standard deviation | | | 2.42 | | |

## A.2. IRMPD/FT-ICR MS of calmodulin


Figure A.1.: Cleavage *c*overage map for one-dimensional top-down and bottom-up anal*y*sis: map of the *c*leavages obtained with the top-down and bottom-up anal*y*sis of CaM using IRMPD FT-ICR MS/MS. The total cleavage coverage for the one-dimensional MS anal*y*sis is ~52%. For comparison, the total cleavage coverage obtained with the 2D MS anal*y*sis is ~40%.

## A.3. ECD 2D MS of calmodulin

Table A.11.: ECD-fragments of the peptide $[91\text{-}106]^{3+}$ ($m/z$ 585.6284).

| **$[91\text{-}106]^{3+}$** ($m/z$ 585.6284) | | | | |
|---|---|---|---|---|
| **Fragment ion** | **Charge** | **Theoretical $m/z$** | **Observed $m/z$** | **Error** |
| ? | 1 | | 342.213739 | |
| $c_3$ | 1 | 379.197596 | 379.197831 | 0.62 |
| $z_3$ | 1 | 401.226885 | 401.227278 | 0.98 |
| ? | 1 | | 487.238880 | |
| $c_4$ | 1 | 507.292559 | 507.293046 | 0.96 |
| $z_5$ | 1 | 543.301113 | 543.300959 | -0.28 |
| ? | 1 | | 574.271357 | |
| $MH_3$ | 3 | 585.628447 | 585.625293 | -5.39 |
| $z_6$ | 1 | 630.333141 | 630.333319 | 0.28 |
| $y_6$ | 1 | 646.351865 | 646.347738 | -6.38 |
| $c_6$ | 1 | 679.340966 | 679.342545 | 2.32 |
| $z_{13}$ | 2 | 689.347786 | 689.348146 | 0.52 |
| $z_7$ | 1 | 743.417205 | 743.417347 | 0.19 |
| $z_{14}$ | 2 | 746.861257 | 746.862370 | 1.49 |
| $c_7$ | 1 | 793.383894 | 793.391204 | 9.21 |
| $z_{15}$ | 2 | 820.395464 | 820.398175 | 3.30 |
| $c_8$ | 1 | 850.405357 | 850.406009 | 0.77 |
| $MH_2$ | 2 | 877.939033 | 877.938987 | -0.05 |
| $z_8$ | 1 | 906.480534 | 906.483194 | 2.93 |

| | | | | |
|---|---|---|---|---|
| $z^{\bullet}_8$ | 1 | 907.488359 | 907.483676 | -5.16 |
| $y_8$ | 1 | 922.499258 | 922.500006 | 0.81 |
| $z_9$ | 1 | 963.501997 | 963.503711 | 1.78 |
| $c_9$ | 1 | 1013.468686 | 1013.469042 | 0.35 |
| $z_{10}$ | 1 | 1077.544925 | 1077.546837 | 1.77 |
| $(a+1)_{10}$ | 1 | 1082.531286 | 1082.541783 | 9.70 |
| $c_{10}$ | 1 | 1126.552750 | 1126.553342 | 0.53 |
| $z_{11}$ | 1 | 1134.566389 | 1134.568732 | 2.06 |
| $z^{\bullet}_{11}$ | 1 | 1135.574214 | 1135.573615 | -0.53 |
| $y_{11}$ | 1 | 1150.585113 | 1150.587706 | 2.25 |
| $c_{11}$ | 1 | 1213.584778 | 1213.586359 | 1.30 |
| $z_{12}$ | 1 | 1249.593332 | 1249.595877 | 2.04 |
| $c_{12}$ | 1 | 1284.621892 | 1284.623288 | 1.09 |
| $c_{13}$ | 1 | 1355.659006 | 1355.659564 | 0.41 |
| $z_{13}$ | 1 | 1377.688295 | 1377.690264 | 1.43 |
| $(a+1)_{14}$ | 1 | 1440.680135 | 1440.690973 | 7.52 |
| $c_{14}$ | 1 | 1484.701599 | 1484.704650 | 2.06 |
| $c_{15}$ | 1 | 1597.785663 | 1597.787929 | 1.42 |
| | | MMA | | 1.24 |
| | | SD | | 3.17 |

Table A.12.: ECD-fragments of the peptide [1-13]$^{2+}$ ($m/z$ 782.3805).

| **[1-13]$^{2+}$** ($m/z$ 782.3805) | | | | |
|---|---|---|---|---|
| **Fragment ion** | **Charge** | **Theoretical $m/z$** | **Observed $m/z$** | **Error [ppm]** |
| $y_2$ | 1 | 294.181218 | 294.181126 | -0.31 |
| $c_3$-$NH_3$ | 1 | 357.140476 | 357.140373 | -0.29 |
| ? | 1 | | 405.213125 | |
| $y_3$ | 1 | 423.223811 | 423.223723 | -0.21 |
| $b_4$?! | 1 | 470.224540 | 470.224472 | -0.15 |
| $z^{\bullet}_4$ | 1 | 479.250026 | 479.249728 | -0.62 |
| $y_4$ | 1 | 494.260925 | 494.261356 | 0.87 |
| $z_5$ | 1 | 591.326265 | 591.326113 | -0.26 |
| $z^{\bullet}_5$ | 1 | 592.334090 | 592.333985 | -0.18 |
| $z_6$ | 1 | 719.384842 | 719.385700 | 1.19 |
| $MH_2$ | 2 | 782.380486 | 782.377468 | -3.86 |
| $z_7$ | 1 | 848.427435 | 848.439665 | 4.42 |
| $z^{\bullet}_7$ | 1 | 849.435260 | 849.439463 | 4.95 |
| $y_7$ | 1 | 864.446160 | 864.438631 | -8.71 |

| | | | | |
|---|---|---|---|---|
| ? | 1 | | 877.936543 | |
| ? | 1 | | 905.452342 | |
| ? | 1 | | 918.458886 | |
| $c_8$ | 1 | 974.442531 | 974.442034 | -0.51 |
| $z_8$ | 1 | 977.470029 | 977.470941 | 0.93 |
| ? | 1 | | 1006.496316 | |
| ? | 1 | | 1069.491702 | |
| $z_9$ | 1 | 1078.517707 | 1078.519106 | 1.30 |
| $y_9$ | 1 | 1094.536431 | 1094.531446 | -4.55 |
| $z_{10}$ | 1 | 1191.601771 | 1191.601933 | 0.14 |
| $y_{10}$ | 1 | 1207.620495 | 1207.621390 | 0.74 |
| ? | 1 | | 1248.631571 | |
| $c^{\bullet}_{11}$ | 1 | 1286.598477 | 1286.598959 | 0.37 |
| $z_{11}$ | 1 | 1319.660349 | 1319.661638 | 0.98 |
| $y_{11}$ | 1 | 1335.679073 | 1335.680865 | 1.34 |
| ? | 1 | | 1390.699966 | |
| $y_{12}$ | 1 | 1450.706016 | 1450.696066 | -6.86 |
| $[MH_2]^{+\bullet}$ | 1 | 1563.753694 | 1563.752307 | -0.89 |
| | | MMA | | -0.42 |
| | | SD | | 2.95 |

Table A.13.: ECD-fragments of the peptide $[106\text{-}125]^{3+}$ ($m/z$ 801.0628).

| **$[106\text{-}125]^{3+}$ ($m/z$ 801.0628)** | | | | |
|---|---|---|---|---|
| **Fragment ion** | **Charge** | **Theoretical $m/z$** | **Observed $m/z$** | **Error** |
| $c^{\bullet}_4$ | 1 | 485.241490 | 485.241968 | 0.99 |
| $z_4$ | 1 | 532.267370 | 532.268018 | 1.22 |
| $z_5$ | 1 | 647.294313 | 647.293998 | -0.49 |
| $MH_3$ | 3 | 801.062775 | 801.059002 | -4.71 |
| $z_7$ | 1 | 875.405320 | 875.406638 | 1.51 |
| $c_{15}$ | 2 | 878.450919 | 878.452613 | 1.93 |
| $z_8$ | 1 | 1004.447913 | 1004.447154 | -0.76 |
| $[MH_3]^{2+\bullet}$ | 2 | 1201.090525 | 1201.088238 | -1.90 |
| $c_{11}$ | 1 | 1283.714019 | 1283.715521 | 1.17 |
| $z_{11}$ | 1 | 1333.606599 | 1333.614112 | 5.63 |
| $c_{12}$ | 1 | 1398.740962 | 1398.742530 | 1.12 |
| $z_{12}$ | 1 | 1503.748512 | 1503.744477 | -2.68 |
| $c_{13}$ | 1 | 1527.783555 | 1527.786308 | 1.80 |
| $z^{\bullet}_{13}$ | 1 | 1633.798930 | 1633.796662 | -1.39 |

| | | | | |
|---|---|---|---|---|
| $c_{14}$ | 1 | 1656.826148 | 1656.823861 | -1.38 |
| $c_{15}$ | 1 | 1755.894562 | 1755.908941 | 8.19 |
| $c_{16}$ | 1 | 1870.921505 | 1870.926154 | 2.48 |
| $z_{16}$ | 1 | 1916.939560 | 1916.951211 | 6.08 |
| $y_{16}$ | 1 | 1932.958284 | 1932.953600 | -2.42 |
| $c_{17}$ | 1 | 1999.964098 | 1999.962744 | -0.68 |
| $c^{\bullet}_{19}$ | 1 | 2243.080822 | 2243.082934 | 0.94 |
| | | MMA | | 0.79 |
| | | SD | | 2.99 |

Table A.14.: ECD-fragments of the peptide [1-30]$^{4+}$ ($m/z$ 848.1622).

| **[1-30]$^{4+}$** ($m/z$ 848.1622) | | | | |
|---|---|---|---|---|
| **Fragment ion** | **Charge** | **Theoretical $m/z$** | **Observed $m/z$** | **Error** |
| $c_{13}$ | 1 | 1562.769679 | 1562.771295 | 1.03 |
| $c_{21}$ | 2 | 1250.615745 | 1250.625452 | 7.76 |
| $c_{23}$ | 2 | 1336.639948 | 1336.644515 | 3.42 |
| $c_{24}$ | 2 | 1394.153420 | 1394.157121 | 2.65 |
| $c_{25}$ | 2 | 1422.664151 | 1422.660860 | -2.31 |
| $z'_{26}$ | 2 | 1453.203109 | 1453.200650 | -1.69 |
| $c_{26}$ | 2 | 1473.187991 | 1473.184443 | -2.41 |
| $z_{27}$ | 2 | 1509.241228 | 1509.240204 | -0.68 |
| $c_{27}$ | 2 | 1529.730023 | 1529.732167 | 1.40 |
| $c_{28}$ | 2 | 1580.253862 | 1580.251289 | -1.63 |
| $z_{29}$ | 2 | 1630.783989 | 1630.785505 | 0.93 |
| $y_{26}$ | 3 | 974.141464 | 974.141041 | -0.43 |
| $y_{27}$ | 3 | 1011.836152 | 1011.832415 | -3.69 |
| $z'_{28}$ | 3 | 1049.518712 | 1049.521571 | 2.72 |
| $z_{29}$ | 3 | 1087.525084 | 1087.525573 | 0.45 |
| $[MH_3]^{2+\bullet}$ | 3 | 1130.547219 | 1130.566550 | 17.10 |
| $MH_4$ | 4 | 848.162233 | 848.159711 | -2.97 |
| | | MMA | | 1.27 |
| | | SD | | 4.82 |

Table A.15.: ECD-fragments of the peptide [90-105]$^{2+}$ ($m/z$ 877.9390).

| **[90-105]$^{2+}$** ($m/z$ 877.9390) | | | | |
|---|---|---|---|---|
| **Fragment ion** | **Charge** | **Theoretical $m/z$** | **Observed $m/z$** | **Error [ppm]** |
| $z^{\bullet}_{7}$ | 1 | 744.425030 | 744.426105 | 1.44 |

| | | | | |
|---|---|---|---|---|
| $MH_2$ | 2 | 877.939033 | 877.934821 | -4.80 |
| $z_8$ | 1 | 906.480534 | 906.479946 | -0.65 |
| $z_9$ | 1 | 963.501997 | 963.506804 | 4.99 |
| $y_9$ | 1 | 979.520721 | 979.520250 | -0.48 |
| $z_{10}$ | 1 | 1077.544925 | 1077.548020 | 2.87 |
| $z^{\bullet}_{10}$ | 1 | 1078.552750 | 1078.555729 | 2.76 |
| ? | 1 | | 1205.604967 | |
| $z_{12}$ | 1 | 1249.593332 | 1249.593885 | 0.44 |
| $z^{\bullet}_{12}$ | 1 | 1250.601157 | 1250.595272 | -4.71 |
| $z_{13}$ | 1 | 1377.688295 | 1377.692423 | 3.00 |
| ? | 1 | | 1448.727590 | |
| $c^{\bullet}_{14}$ | 1 | 1483.693774 | 1483.694722 | 0.64 |
| $z_{14}$ | 1 | 1492.715238 | 1492.714919 | -0.21 |
| $c^{\bullet}_{15}$ | 1 | 1596.777838 | 1596.779176 | 0.84 |
| $z^{\bullet}_{15}$ | 1 | 1640.791477 | 1640.769312 | -13.51 |
| $[MH_2]^{+\bullet}$ | 1 | 1754.870790 | 1754.869293 | -0.85 |
| | | MMA | | -0.55 |
| | | SD | | 4.30 |

Table A.16.: ECD-fragments of the peptide $[14\text{-}30]^{2+}$ ($m/z$ 922.9493).

| **$[14\text{-}30]^{2+}$ ($m/z$ 922.9493)** | | | | |
|---|---|---|---|---|
| **Fragment ion** | **Charge** | **Theoretical $m/z$** | **Observed $m/z$** | **Error [ppm]** |
| $y_6$ | 1 | 620.361367 | 620.361292 | -0.12 |
| $MH_2$ | 2 | 922.949263 | 922.945412 | -4.17 |
| $c^{\bullet}_{8}$ | 1 | 954.480534 | 954.483222 | 2.82 |
| $z^{\bullet}_{10}$ | 1 | 1020.520781 | 1020.523003 | 2.18 |
| $c^{\bullet}_{10}$ | 1 | 1126.528940 | 1126.527902 | -0.92 |
| $z_{11}$ | 1 | 1134.539899 | 1134.537185 | -2.39 |
| $c^{\bullet}_{11}$ | 1 | 1241.555883 | 1241.555472 | -0.33 |
| $z_{12}$ | 1 | 1281.608313 | 1281.615907 | 5.93 |
| $c^{\bullet}_{12}$ | 1 | 1298.577347 | 1298.578352 | 0.77 |
| $z_{13}$ | 1 | 1394.692377 | 1394.693966 | 1.14 |
| $c^{\bullet}_{13}$ | 1 | 1399.625026 | 1399.624306 | -0.51 |
| $z_{14}$ | 1 | 1481.724405 | 1481.727976 | 2.41 |
| $y_{14}$ | 1 | 1497.743129 | 1497.744202 | 0.72 |
| $c^{\bullet}_{14}$ | 1 | 1512.709090 | 1512.706547 | -1.68 |
| $c_{15}$ | 1 | 1614.764593 | 1614.768230 | 2.25 |
| $z^{\bullet}_{15}$ | 1 | 1629.800644 | 1629.799363 | -0.79 |

| | | | | |
|---|---|---|---|---|
| ? | 1 | | 1653.779388 | |
| $c_{16}$ | 1 | 1715.812272 | 1715.813250 | 0.57 |
| $[MH_2]^{+\bullet}$ | 1 | 1844.891250 | 1844.887701 | -1.92 |
| | | MMA | | 0.33 |
| | | SD | | 2.25 |

Table A.17.: ECD-fragments of the peptide $[127\text{-}148]^{2+}$ ($m/z$ 1245.5436).

| **$[127\text{-}148]^{2+}$** ($m/z$ 1245.5436) | | | | |
|---|---|---|---|---|
| **Fragment ion** | **Charge** | **Theoretical $m/z$** | **Observed $m/z$** | **Error [ppm]** |
| $MH_2$ | 2.00 | 1245.543605 | 1245.531140 | -10.01 |
| $z^{\bullet}{}_{20}$ | 1.00 | 2274.989327 | 2274.986044 | -1.44 |
| $z'_{21}$ | 1.00 | 2346.026441 | 2346.009311 | -7.30 |
| $y_{21}$ | 1.00 | 2361.037340 | 2361.011012 | -11.15 |
| $[MH_2]^{+\bullet}$ | 1.00 | 2491.079933 | 2491.086190 | 2.51 |
| | | MMA | | 0.33 |
| | | SD | | 2.25 |

Table A.9.: Peak assignment for bottom-up 1D IRMPD FT-ICR MS Calmodulin.

| **Peptide** | **Theoretical** $m/z$ | **Observed** $m/z$ | **Error [ppm]** | **Peptide sequence** | **Charge** |
|---|---|---|---|---|---|
| 31_37 | 403.215464 | 403.21421 | -3.11 | ELGTVMR | 2 |
| 75_77 | 406.2483 | 406.24719 | -2.73 | KMK | 1 |
| 76_86 | 451.538131 | 451.53762 | -1.13 | MKDTDSEEEIR | 3 |
| 14_21 | 478.739827 | 478.73947 | -0.75 | EAFSLFDK | 2 |
| 75_86 | 494.236452 | 494.23584 | -1.24 | KMKDTDSEEEIR | 3 |
| 87_90 | 522.267074 | 522.26659 | -0.93 | EAFR | 1 |
| 78_90 | 532.909393 | 532.90945 | 0.11 | DTDSEEEIREAFR | 3 |
| 78_86 | 547.235835 | 547.23583 | -0.01 | DTDSEEEIR | 2 |
| 91_106 | 585.628449 | 585.62840 | -0.08 | VFDKDGNGYISAAELR | 3 |
| d91_106 | 585.956454 | 585.95641 | -0.08 | VFDKDGNGYISAAELR | 3 |
| 14_30 | 615.6352697 | 615.63524 | -0.05 | EAFSLFDKDGDGTITTK | 3 |
| 76_86 | 676.803559 | 676.80370 | 0.21 | MKDTDSEEEIR | 2 |
| 1_13 | 782.380488 | 782.38055 | 0.08 | Ac-ADQLTEEQIAEFK | 2 |
| 78_90 | 798.8604515 | 798.86131 | 1.07 | DTDSEEEIREAFR | 2 |
| 106_125 | 801.0627777 | 801.06433 | 1.94 | HVMTNLGEKLTDEEVDEMIR | 3 |
| 31_37 | 805.423652 | 805.42358 | -0.09 | ELGTVMR | 1 |
| 91_106 | 877.9390355 | 877.94069 | 1.88 | VFDKDGNGYISAAELR | 2 |
| d91_106 | 878.431044 | 878.43151 | 0.53 | VFDKDGNGYISAAELR | 2 |
| 22_30 | 907.436721 | 907.43861 | 2.08 | DGDGTITTK | 1 |
| 14_30 | 922.949267 | 922.94928 | 0.01 | EAFSLFDKDGDGTITTK | 2 |
| 14_21 | 956.472377 | 956.47375 | 1.44 | EAFSLFDK | 1 |
| 78_86 | 1093.464394 | 1093.46538 | 0.90 | DTDSEEEIR | 1 |
| 127_148 | 1245.543609 | 1245.54480 | 0.96 | EADIDGDGQVNYEEFVQMMTAK | 2 |
| 1_13 | 1563.7537 | 1563.75544 | 1.11 | Ac-ADQLTEEQIAEFK | 1 |
| Average Error | | | 0.09 | | |
| Standard deviation | | | 1.28 | | |

Table A.10.: Peak assignments to the auto*c*orrelation line of the ECD 2D MS spe*c*trum of tr*y*pti*c*-digested CaM. For the precursor ions indicated with a * s*y*mbol, MS/MS anal*y*sis is available in tables A.11 to A.17 below.

| **Pre*c*ursor ion** | **Sequen*c*e** | **Charge** | **Theoreti*c*al** $m/z$ | **Observed** $m/z$ | **Error [ppm]** |
|---|---|---|---|---|---|
| [31-37] | ELGTVMR | 2 | 403.215464 | 403.215958 | -1.22 |
| ? | ? | 2 | | 496.739960 | |
| [78-90] | DTDSEEEIREAFR | 3 | 532.909393 | 532.909029 | 0.68 |
| [87-106] | EAFRVFDKDGNGYISAAELR | 4 | 565.285463 | 565.283866 | 2.83 |
| [91-106]* | VFDKDGNGYISAAELR | 3 | 585.628449 | 585.626130 | 3.96 |
| ? | ? | 4 | | 601.050073 | |
| [14_30] | EAFSLFDKDGDGTITTK | 3 | 615.635270 | 615.635413 | -0.23 |
| [75-90] | KMKDTDSEEEIREAFR | 3 | 661.9861946 | 661.985819 | 0.57 |
| [76-86] | MKDTDSEEEIR | 2 | 676.803559 | 676.802803 | 1.12 |
| ? | ? | 3 | | 695.720320 | |
| ? | ? | 3 | | 705.062640 | |
| [75-86] | KMKDTDSEEEIR | 2 | 740.851038 | 740.850213 | 1.11 |
| [1-13]* | ADQLTEEQIAEFK | 2 | 782.380488 | 782.380361 | 0.16 |
| [106-125]* | HVMTNLGE**K**LTDEEVDEMIR | 3 | 801.062778 | 801.062206 | 0.71 |
| [127-148]* | EADIDGDGQVNYEEFVQMMTAK | 3 | 830.6981616 | 830.697353 | 0.97 |
| [1-30]* | ADQLTEEQIAEFKEAFSLFDKDGDGTITTK | 4 | 848.162233 | 848.163606 | -1.62 |
| [90-105] | VFDKDGNGYISAAELR | 2 | 877.939036 | 877.939213 | -0.20 |
| [14-30] | EAFSLFDKDGDGTITTK | 2 | 922.949267 | 922.950608 | -1.45 |
| [1-37] | DTDSEEEIREAFRVFDK | 4 | 1044.763689 | 1044.768953 | -5.04 |
| [1-30] | ADQLTEEQIAEFKEAFSLFDKDGDGTITTK | 3 | 1130.547219 | 1130.555453 | -7.28 |
| [107-148] | HVMTNLGE**K**LTDEEVDEMIREADIDGDGQVNYEEFVQMMTAK | 4 | 1218.814424 | 1218.820736 | -5.18 |
| [127-148] | EADIDGDGQVNYEEFVQMMTAK | 2 | 1245.543609 | 1245.543275 | 0.27 |
| [38-74] | SLGQNPTEAELQDMINEVDADGNGTIDFPEFLTMMAR | 3 | 1357.2869750 | 1357.300378 | -9.87 |
| | | | MMA | | -1.04 |
| | | | SD | | 3.38 |

# B. Supporting Information for Chapter 4

# B.1. Mass spectrometry of Ubiquitin

## B.1.1. One-dimensional MS of Ubiquitin

Table B.1.: Fragments assigned to the one-dimensional CAD MS/MS spectrum of Ubiquitin in denaturing conditions.

| **Ion** | **Observed m/z** | **Charge** | **Theoretical m/z** | **Error [ppm]** |
|---|---|---|---|---|
| $b_3$ | 373.190403 | 1 | 373.190403 | 0.00 |
| ${y_{24}}^{5+}$ | 546.10686 | 5 | 546.107026 | -0.30 |
| $[b_{15}\text{-}H_2O]^{3+}$ | 553.327318 | 3 | 553.328505 | -2.15 |
| $PSDTIENVKAKIQDK^{3+}$ | 556.637048 | 3 | 556.636944 | 0.19 |
| ${b_{15}}^{3+}$ | 559.331591 | 3 | 559.332027 | -0.78 |
| $[y_{15}\text{-}H_2O]^{3+}$ | 563.667944 | 3 | 563.668801 | -1.52 |
| TLEVE | 572.292488 | 1 | 572.292621 | -0.23 |
| $(a_{16}\text{-}1)^{3+}$ | 587.01024 | 3 | 587.005767 | 7.62 |
| $b_{16}\text{-}2H_2O^{3+}$ | 590.338617 | 3 | 590.339182 | -0.96 |
| ${a_{16}}^{3+}$ | 593.015934 | 3 | 593.014600 | 2.25 |
| $b_{16}\text{-}H_2O^{3+}$ | 596.342207 | 3 | 596.342703 | -0.83 |
| ${b_{16}}^{3+}$ | 602.345959 | 3 | 602.346225 | -0.44 |
| ${y_{37}}^{7+}$ | 608.480702 | 7 | 608.481459 | -1.24 |
| QDKEGIPPDQQRLIFA/ PPDQQRLIFAGKQLED$^{3+}$ | 612.991031 | 3 | 612.991610 | -0.94 |
| GKTITL/LTGKTI | 614.387142 | 1 | 614.387189 | -0.08 |
| $b_5$ | 619.32719 | 1 | 619.327231 | -0.07 |
| ${a_{17}}^{3+}$ | 626.038234 | 3 | 626.037400 | 1.33 |
| $[b_{17}\text{-}H_2O]^{3+}$ | 629.365018 | 3 | 629.365508 | -0.78 |
| ${b_{17}}^{3+}$ | 635.368997 | 3 | 635.369030 | -0.05 |
| ? | 637.304731 | 1 | | |
| $y_{28}\text{-}H_2O^{5+}$ | 639.54739 | 5 | 639.547349 | 0.06 |
| $y_{57}\text{-}H_2O^{10+}$ | 642.245561 | 10 | 642.2498312 | -6.65 |
| $y_{58}\text{-}H_2O^{10+}$ | 651.952414 | 10 | 651.9551076 | -4.13 |
| ${y_{58}}^{10+}$ | 653.755757 | 10 | 653.756164 | -0.62 |
| SDTIENVKAKIQDKEGIP/ PSDTIENVKAKIQDKEGI$^{3+}$ | 656.352721 | 3 | 656.352985 | -0.40 |
| $y_{59}\text{-}H_2O^{10+}$ | 664.857482 | 10 | 664.859367 | -2.84 |
| $b_{18}\text{-}H_2O^{3+}$ | 672.379089 | 3 | 672.379706 | -0.92 |
| ${b_{18}}^{3+}$ | 678.382672 | 3 | 678.383228 | -0.82 |

| | | | | |
|---|---|---|---|---|
| $y_{24}{}^{4+}$ | 682.381965 | 4 | 682.381964 | 0.00 |
| $y_{37}{}^{6+}$ | 709.726161 | 6 | 709.727156 | -1.40 |
| $?^{12+}$ | 711.388751 | 12 | | |
| MH-$H_2O^{12+}$ | 712.807595 | 12 | | |
| $y_{57}$-$NH_3{}^{9+}$ | 713.610142 | 9 | 713.608339 | 2.53 |
| $y_{57}{}^{9+}$ | 715.50125 | 9 | 715.500178 | 1.50 |
| $?^{12+}$ | 717.553521 | 12 | | |
| $y_{58}$-$H_2O^{9+}$ | 724.281825 | 9 | 724.282644 | -1.13 |
| $y_{58}{}^{9+}$ | 726.286725 | 9 | 726.283818 | 4.00 |
| $y_{59}$-$H_2O^{9+}$ | 738.619766 | 9 | 738.620710 | -1.28 |
| $y_{74}{}^{11+}$ | 755.600089 | 11 | 755.599792 | 0.39 |
| PSDTIENVKAKIQDKEGIPPD$^{3+}$ | 759.397666 | 3 | 759.397142 | 0.69 |
| PSDTIENVKAKIQDKEGIPP-DQQRLIFAGKQLED/ EVEPSDTIENVKAKIQDKE-GIPPDQQRLIFAGKQ$^{5+}$ | 761.405389 | 5 | 761.405302 | 0.11 |
| IFVKTLTGKTITLE$^{2+}$ | 773.465138 | 2 | 773.466168 | -1.33 |
| $y_{55}{}^{8+}$ | 779.555314 | 8 | 779.554419 | 1.15 |
| $b_{14}{}^{2+}$ | 781.954531 | 2 | 781.952371 | 2.76 |
| [AKIQDKEGIPPDQQRLIFAGKQ-LEDGRTLSDYNIQ -2$H_2O$-$NH_3$]$^{5+}$ | 783.609567 | 5 | 783.609728 | -0.21 |
| $y_{56}{}^{8+}$ | 793.932888 | 8 | 793.932787 | 0.13 |
| [$y_{57}$-$NH_3$]$^{8+}$ | 802.687294 | 8 | 802.683472 | 4.76 |
| $y_{57}{}^{8+}$ | 804.808837 | 8 | 804.811791 | -3.67 |
| [$y_{58}$-$H_2O$]$^{8+}$ | 814.692368 | 8 | 814.692066 | 0.37 |
| $y_{58}{}^{8+}$ | 816.944851 | 8 | 816.943386 | 1.79 |
| $y_{59}$-$H_2O^{8+}$ | 830.822849 | 8 | 830.822390 | 0.55 |
| $b_{15}{}^{2+}$ | 838.493595 | 2 | 838.494403 | -0.96 |
| TGKTITLE | 844.476915 | 1 | 844.477462 | -0.65 |
| [$b_{16}$-$H_2O$]$^{2+}$ | 894.011485 | 2 | 894.010417 | 1.19 |
| $b_{16}{}^{2+}$ | 903.015696 | 2 | 903.015700 | 0.00 |
| [KTLTGKTITLEVEPSDTIENVKA-KIQDKEGIPPDQQRLIFAGKQL-EDGRTLSDYNIQK -3$NH_3$]$^{7+}$ | 921.061173 | 7 | 921.060924 | 0.27 |
| [$b_{17}$-$H_2O$]$^{2+}$ | 943.543687 | 2 | 943.544624 | -0.99 |
| TLSDYNIQKESTLHLVL-3$H_2O^{2+}$ | 951.504578 | 2 | 951.5096275 | -5.31 |
| $b_{17}{}^{2+}$ | 952.549776 | 2 | 952.549907 | -0.14 |
| $b_{18}{}^{2+}$ | 1017.071207 | 2 | 1017.071204 | 0.00 |

| | | | | |
|---|---|---|---|---|
| $b_9$ | 1062.601253 | 1 | 1062.601616 | -0.34 |
| TGKTITLEVE | 1072.592385 | 1 | 1072.588470 | 3.65 |

Table B.2.: Fragments assigned to the one-dimensional CAD/IRMPD MS/MS/MS spectrum of Ubiquitin in denaturing conditions.

| **Ion** | **Charge** | **Theoretical m/z** | **Observed m/z** | **Error [ppm]** |
|---|---|---|---|---|
| ${y_{24}}^{4+}$ | 4 | 682.381964 | 682.381962 | 0.00 |
| ${y_{18}}^{3+}$ | 3 | 699.735763 | 699.735768 | 0.01 |
| ${y_{13}}^{2+}$ | 2 | 725.928076 | 725.927906 | -0.23 |
| ${y_{58}}^{9+}$ | 9 | 726.283818 | 726.283984 | 0.23 |
| PSDTIENVKAKIQDKEG-IPPDQQRLIFAGKQLED | 5 | 761.405302 | 761.405594 | 0.38 |

Table B.3.: Fragments assigned to the one-dimensional CAD/ECD MS/MS/MS spectrum of Ubiquitin in denaturing conditions.

| **Fragment Ion** | **Charge** | **Theoretical m/z** | **Observed m/z** | **Error [ppm]** |
|---|---|---|---|---|
| $z_{11}$ | 2 | 609.881402 | 609.881429 | 0.04 |
| ? | 2 | | 625.365888 | |
| $z_{17}$ | 3 | 640.041745 | 640.041696 | -0.08 |
| $c_{12}$ | 2 | 657.374811 | 657.375106 | 0.45 |
| $c_6$ | 1 | 660.319899 | 660.319906 | 0.01 |
| $c_{13}$ | 2 | 721.404100 | 721.404129 | 0.04 |
| ? | 2 | | 725.418597 | |
| ? | 5 | | 733.002565 | |
| $c_{26}$ | 4 | 733.652912 | 733.652717 | -0.27 |
| $c_{33}$ | 5 | 741.805223 | 741.804808 | -0.56 |
| $c_{41}$ | | 769.572826 | 769.573480 | 0.85 |
| $c_7$ | 1 | 774.362827 | 774.362130 | -0.90 |
| | 8 | | 811.439598 | |
| | 8 | | 815.066907 | |
| $y_{58}8+$● | 8 | | 817.068801 | |
| ? | 3 | | 838.465293 | |
| $(z+1)_{52}$ | 7 | 839.606120 | 839.603881 | -2.67 |
| $c_{15}$ | 2 | 842.965053 | 842.964140 | -1.08 |
| ${(z+1)_{45}}^{6+}$ | | | 849.128221 | |

| | | | | |
|---|---|---|---|---|
| $c_{23}$ | 3 | 850.445044 | 850.442753 | -2.69 |
| ? | 6 | | 851.793043 | |
| $(z+1)_{53}$ | 7 | 858.040776 | 858.038579 | -2.56 |
| $(z+1)_{46}$ | | 870.468853 | 870.466455 | -2.75 |
| $c_{46}$ | 6 | 871.626680 | 871.625951 | -0.84 |
| $(z+2)_{54}$ | 7 | 874.339618 | 874.337774 | -2.11 |
| $c_{54}$ | 7 | 878.618242 | 878.618386 | 0.16 |
| $z_{23}$ | 3 | 885.160131 | 885.158901 | -1.39 |
| $c_{47}$ | 6 | 886.132018 | 886.131656 | -0.41 |
| $(z+2)_{47}$ | 6 | 889.483400 | 889.480821 | -2.90 |
| $c_{40}$ | 5 | 890.673270 | 890.672660 | -0.68 |
| ? | 7 | | 895.060831 | |
| ? | 5 | | 900.734911 | |
| $c_{24}$ | 3 | 902.478748 | 902.476943 | -2.00 |
| $(z+2)_{56}$ | 7 | 905.207421 | 905.205357 | -1.59 |
| ? | 5 | | 916.252058 | |
| ? | 7 | | 917.496135 | |
| $c_{41}$ | 5 | 923.285936 | 923.283356 | -2.79 |
| $c_{57}$ | 7 | 925.219190 | 925.214151 | -5.45 |
| | 7 | | 927.503526 | |
| | 7 | | 933.645723 | |
| $z_{25}$ | 3 | 942.509600 | 942.509592 | -0.01 |
| ? | 3 | | 947.846968 | |
| $z_{17}$ | 2 | 959.558980 | 959.556948 | -2.12 |
| $(z+2)_{52}$ | 6 | 979.539800 | 979.536826 | -3.04 |
| $c_{53}$ | 6 | 998.869885 | 998.867055 | -2.83 |
| $c_{36}$ | 4 | 1009.042089 | 1009.039507 | -2.56 |
| ? | 3 | | 1066.239463 | |
| $c_{20}$ | 2 | 1089.591878 | 1089.589662 | -2.03 |

### B.1.2. MS/2D MS of Ubiquitin

Table B.4.: CAD fragments assigned to the autocorrelation line of the CAD/IRMPD MS/2D MS experiment.

| **Percursor Ion** | **Charge** | **Theoretical m/z** | **Observed m/z** | **Error [ppm]** |
|---|---|---|---|---|
| $b_3$ | 1 | 373.190403 | 373.191394 | 2.66 |
| ${a_8}^{2+}$ | 2 | 467.283100 | 467.283281 | 0.39 |

| | | | | |
|---|---|---|---|---|
| $y_{18}{}^{4+}$ | 4 | 525.053641 | 525.053361 | -0.53 |
| $y_{24}$-$H_2O^{5+}$ | 5 | 542.504913 | 542.505521 | 1.12 |
| $y_{24}{}^{5+}$ | 5 | 546.107026 | 546.106943 | -0.15 |
| $b_{15}$-$H_2O^{3+}$ | 3 | 553.328505 | 553.328415 | -0.16 |
| $b_{15}{}^{3+}$ | 3 | 559.332027 | 559.331346 | -1.22 |
| TLEVE | 1 | 572.292621 | 572.293434 | 1.42 |
| $y_{16}$-$3NH_3{}^{3+}$ | 3 | 590.340462 | 590.339264 | -2.03 |
| $IFVKTLTGKTITLEVE^{3+}$ | 3 | 592.016873 | 592.016392 | -0.81 |
| $a_{16}{}^{3+}$ | 3 | 593.014600 | 593.014908 | 0.52 |
| $b_{16}$-$H_2O^{3+}$ | 3 | 596.342703 | 596.342854 | 0.25 |
| $b_{16}{}^{3+}$ | 3 | 602.346225 | 602.346710 | 0.81 |
| $y_{16}{}^{3+}$ | 3 | 607.367011 | 607.365189 | -3.00 |
| $y_{37}{}^{7+}$ | 7 | 608.481459 | 608.481054 | -0.67 |
| QDKEGIPPDQQRLIFA/ PPDQQRLIFAGKQLED | 3 | 612.991610 | 612.991990 | 0.62 |
| $b_5$ | 1 | 619.327231 | 619.327581 | 0.56 |
| $a_{17}{}^{3+}$ | 3 | 626.037400 | 626.037114 | -0.46 |
| $b_{17}$-$H_2O^{3+}$ | 3 | 629.365508 | 629.365523 | 0.02 |
| $b_{17}{}^{3+}$ | 3 | 635.369030 | 635.368609 | -0.66 |
| $y_{28}$-$H_2O^{5+}$ | 5 | 639.547349 | 639.546968 | -0.60 |
| $y_{57}$-$NH_3{}^{10+}$ | 10 | 642.348233 | 642.352263 | 6.27 |
| $y_{28}{}^{5+}$ | 5 | 643.149462 | 643.150526 | 1.65 |
| $y_{58}$-$H_2O^{10+}$ | 10 | 651.955108 | 651.952900 | -3.39 |
| $y_{40}{}^{7+}$ | 7 | 652.643241 | 652.643684 | 0.68 |
| $y_{58}{}^{10+}$ | 10 | 653.756164 | 653.756672 | 0.78 |
| $y_{59}$-$2H_2O^{10+}$ | 10 | 663.058311 | 663.058045 | -0.40 |
| $y_{59}$-$H_2O^{10+}$ | 10 | 664.859367 | 664.858883 | -0.73 |
| $y_{59}{}^{10+}$ | 10 | 666.660424 | 666.662005 | 2.37 |
| $a_{18}{}^{3+}$ | 3 | 669.051600 | 669.051691 | 0.14 |
| $b_{18}$-$H_2O^{3+}$ | 3 | 672.379706 | 672.379846 | 0.21 |
| $y_{60}$-$H_2O^{10+}$ | 10 | 674.766208 | 674.767670 | 2.17 |
| $y_{60}{}^{10+}$ | 10 | 676.567265 | 676.566956 | -0.46 |
| $b_{18}{}^{3+}$ | 3 | 678.383228 | 678.382980 | -0.37 |
| $y_{54}{}^{9+}$ | 9 | 681.821661 | 681.822405 | 1.09 |
| $y_{24}{}^{4+}$ | 4 | 682.381964 | 682.379414 | -3.74 |
| $y_{61}$-$H_2O$-$NH_3{}^{10+}$ | 10 | 685.967813 | 685.969337 | 2.22 |
| $y_{61}$-$H_2O^{10+}$ | 10 | 687.670468 | 687.670219 | -0.36 |
| $y_{55}{}^{9+}$ | 9 | 693.049181 | 693.049007 | -0.25 |

| | | | | |
|---|---|---|---|---|
| $y_{49}{}^{8+}$ | 8 | 694.006839 | 694.001077 | -8.30 |
| $y_{31}{}^{5+}$ | 5 | 694.380171 | 694.381104 | 1.34 |
| QIFVKT-$H_2O$ | 1 | 699.418800 | 699.407423 | -16.27 |
| $y_{62}{}^{10+}$ | 10 | 700.779931 | 700.781467 | 2.19 |
| $y_{56}{}^{9+}$ | 9 | 705.829952 | 705.827151 | -3.97 |
| $y_{37}{}^{6+}$ | 6 | 709.727156 | 709.727736 | 0.82 |
| $y_{58}$-$H_2O^{9+}$ | 9 | 724.282644 | 724.282044 | -0.83 |
| $y_{58}{}^{9+}$ | 9 | 726.283818 | 726.283741 | -0.11 |
| $y_{59}$-$2H_2O^{9+}$ | 9 | 736.619537 | 736.619133 | -0.55 |
| $y_{59}$-$H_2O^{9+}$ | 9 | 738.620710 | 738.619647 | -1.44 |
| $y_{59}{}^{9+}$ | 9 | 740.621884 | 740.623576 | 2.28 |
| $y_73^{11+}$ | 11 | 745.319422 | 745.320336 | 1.23 |
| $b_6$ | 1 | 747.422194 | 747.424171 | 2.65 |
| $b_{60}$-$H_2O^{9+}$ | 9 | 747.954717 | 747.960731 | 8.04 |
| $y_{60}{}^{9+}$ | 9 | 751.629486 | 751.630817 | 1.77 |
| $y_{53}{}^{8+}$ | 8 | 752.787951 | 752.782861 | -6.76 |
| $y_{74}$-$H_2O^{11+}$ | 11 | 753.962468 | 753.960516 | -2.59 |
| $y_{74}{}^{11+}$ | 11 | 755.599792 | 755.599866 | 0.10 |
| PSDTIENVKAKIQDKEGIP-PDQQRLIFAGKQLED-$H_2O$/ EVEPSDTIENVKAKIQDKE-GIPPDQQRLIFAGKQ-$H_2O$ | 5 | 757.803189 | 757.803361 | 0.23 |
| PSDTIENVKAKIQDKEGIPPD$^{3+}$ | 3 | 759.397142 | 759.395251 | -2.49 |
| PSDTIENVKAKIQDKEG-IPPDQQRLIFAGKQLED/ EVEPSDTIENVKAKIQD-KEGIPPDQQRLIFAGKQ | 5 | 761.405302 | 761.403606 | -2.23 |
| $y_{54}{}^{8+}$ | 8 | 766.923459 | 766.927642 | 5.45 |
| $b_{14}$-$H_2O^{2+}$ | 2 | 772.947088 | 772.946933 | -0.20 |
| $b_{14}$-$NH_3{}^{2+}$ | 2 | 773.439096 | 773.447542 | 10.92 |
| $y_{55}{}^{8+}$ | 8 | 779.554419 | 779.554794 | 0.48 |
| $b_{14}{}^{2+}$ | 2 | 781.952371 | 781.952869 | 0.64 |
| AKIQDKEGIPPDQQRLIFAGKQ-LEDGRTLSDYNIQ-$2H_2O$-$NH_3{}^{5+}$ | 5 | 783.609728 | 783.610949 | 1.56 |
| PSDTIENVKAKIQDKEGIPPD-QQRLIFAGKQLEDGRTLSDYN$^{6+}$ | 6 | 785.742222 | 785.744697 | 3.15 |
| $y_{56}{}^{8+}$ | 8 | 793.932787 | 793.932648 | -0.18 |
| $y_{57}$-$H_2O$-$NH_3{}^{8+}$ | 8 | 800.432151 | 800.424902 | -9.06 |

| | | | | |
|---|---|---|---|---|
| $y_{57}$-$H_2O^{8+}$ | 8 | 802.560470 | 802.560512 | 0.05 |
| ${y_{28}}^{4+}$ | 4 | 803.685009 | 803.687056 | 2.55 |
| ${y_{57}}^{8+}$ | 8 | 804.811791 | 804.811523 | -0.33 |
| LEVEPSDTIENVKAKIQDKEGIP-PDQQRLIFAGKQLEDGRTLSDY-NIQKESTLHLVL | 8 | 806.055622 | 806.060342 | 5.86 |
| ${y_{43}}^{6+}$ | 6 | 811.107255 | 811.106321 | -1.15 |
| $y_{58}$-$2H_2O^{8+}$ | 8 | 812.440745 | 812.444861 | 5.07 |
| $y_{58}$-$H_2O$-$NH_3{}^{8+}$ | 8 | 812.563747 | 812.566680 | 3.61 |
| $y_{58}$-$H_2O^{8+}$ | 8 | 814.692066 | 814.692955 | 1.09 |
| ${y_{58}}^{8+}$ | 8 | 816.943386 | 816.943131 | -0.31 |
| IFVKTLTGKTITLEV$^{2+}$ | 2 | 823.000375 | 822.999583 | -0.96 |
| $y_{59}$-$H_2O^{8+}$ | 8 | 830.822390 | 830.821107 | -1.54 |
| $b_{52}$-$H_2O^{7+}$ | 7 | 831.879764 | 831.880029 | 0.32 |
| ${b_{52}}^{7+}$ | 7 | 834.452702 | 834.451264 | -1.72 |
| ${b_{15}}^{2+}$ | 2 | 838.494403 | 838.495668 | 1.51 |
| ${y_{37}}^{5+}$ | 5 | 851.471132 | 851.471395 | 0.31 |
| ${y_{15}}^{2+}$ | 2 | 854.004846 | 854.005053 | 0.24 |
| ${b_{39}}^{5+}$ | 5 | 862.466767 | 862.467191 | 0.49 |
| TIENVKAKIQDKEGIP-$H_2O$-$NH_3{}^{2+}$ | 2 | 865.477796 | 865.487341 | 11.03 |
| $y_{16}$-$3NH_3{}^{2+}$ | 2 | 885.007055 | 885.007195 | 0.16 |
| $b_{16}$-$H_2O^{2+}$ | 2 | 894.010417 | 894.010831 | 0.46 |
| ${b_{16}}^{2+}$ | 2 | 903.015700 | 903.015258 | -0.49 |
| ${y_{40}}^{5+}$ | 5 | 913.297627 | 913.299265 | 1.79 |
| KTLTGKTITLEVEPSDTIEN-VKAKIQDKEGIPPDQQRL-IFAGKQLEDGRTLSDYNI-QK-$3NH_3{}^{7+}$ | 7 | 921.060924 | 921.066964 | 6.56 |
| $b_{17}$-$H_2O^{2+}$ | 2 | 943.544624 | 943.546721 | 2.22 |
| PSDTIENVKAKIQDKEGIP-PDQQRLIFAGKQLED-$H_2O$ EVEPSDTIENVKAKIQDKE-GIPPDQQRLIFAGKQ-$H_2O$ | 4 | 947.002167 | 947.004641 | 2.61 |
| ${b_{17}}^{2+}$ | 2 | 952.549907 | 952.551213 | 1.37 |
| ${y_{17}}^{2+}$ | 2 | 967.568342 | 967.567515 | -0.85 |
| GKTITLEVE | 1 | 971.540791 | 971.539681 | -1.14 |
| EPSDTIENVKAKIQDKEG-IPPDQQRLIFAGKQLED$^{4+}$ | 4 | 983.765457 | 983.770005 | 4.62 |

| | | | | |
|---|---|---|---|---|
| $b_{18}$-2$H_2O^{2+}$ | 2 | 999.060639 | 999.060657 | 0.02 |
| $b_{18}$-$H_2O^{2+}$ | 2 | 1008.065921 | 1008.069347 | 3.40 |
| $b_{18}{}^{2+}$ | 2 | 1017.071204 | 1017.072166 | 0.95 |
| TGKTITLEVE | 1 | 1072.588470 | 1072.585238 | -3.01 |

Table B.5.: IRMPD-fragments assigned to the extraction of the horizontal scan of the CAD-fragment $y_{24}{}^{5+}$.

| **Fragment ion** | **Charge** | **Theoretical m/z** | **Observed m/z** | **Error [ppm]** |
|---|---|---|---|---|
| $y_{12}{}^{3+}$ | 3 | 441.273611 | 441.285539 | 27.03 |
| $y_{18}{}^{4+}$ | 4 | 525.053641 | 525.058609 | 9.46 |
| [$y_{24}$-2$H_2O$]$^{5+}$ | 5 | 538.902800 | 538.906146 | 6.21 |
| [$y_{24}$-$H_2O$]$^{5+}$ | 5 | 542.504913 | 542.508050 | 5.78 |
| GRTLSD | 1 | 630.320567 | 630.320067 | -0.79 |

Table B.6.: IRMPD-fragments assigned to the extraction of the horizontal scan of the CAD-fragment $b_{16}{}^{3+}$.

| **Fragment ion** | **Charge** | **Theoretical m/z** | **Observed m/z** | **Error [ppm]** |
|---|---|---|---|---|
| $b_{16}$-$H_2O^{3+}$ | 3 | 596.342703 | 596.343519 | 1.37 |
| $b_{14}$-$H_2O$-$NH_3{}^{2+}$ | 2 | 764.433814 | 764.462470 | 37.49 |

Table B.7.: IRMPD-fragments assigned to the extraction of the horizontal scan of the CAD-fragment $y_{58}10+$.

| **Fragment ion** | **Charge** | **Theoretical *m/z*** | **Observed *m/z*** | **Error [ppm]** |
|---|---|---|---|---|
| $y_{18}{}^{4+}$ | 4 | 525.053641 | 525.058313 | 8.90 |
| $y_{24}$-2$H_2O^{5+}$ | 5 | 538.902800 | 538.906666 | 7.17 |
| $y_{24}{}^{5+}$ | 5 | 546.107026 | 546.110077 | 5.59 |
| QDKEGIPPDQQRLIFA/ PPDQQRLIFAGKQLED$^{3+}$ | 3 | 612.991610 | 612.991095 | -0.84 |
| VKAKIQDKEGIPPDQQRLI-FAGKQLEDGRTLSDY-2$H_2O$ | 5 | 761.410599 | 761.405127 | -7.19 |

Table B.8.: IRMPD-fragments assigned to the extraction of the horizontal scan of the CAD-fragment $b_{17}{}^{3+}$.

| **Fragment ion** | **Charge** | **Theoretical** ***m/z*** | **Observed** ***m/z*** | **Error [ppm]** |
|---|---|---|---|---|
| $b_{17}$-$H_2O^{3+}$ | 3 | 629.365508 | 629.365431 | -0.12 |
| IFVKTLTGKTITLEV-$H_2O$ | 2 | 813.995093 | 813.994818 | -0.34 |
| IFVKTLTGKTITLEV | 2 | 823.000375 | 823.005077 | 5.71 |

Table B.9.: IRMPD-fragments assigned to the extraction of the horizontal scan of the CAD-fragment $b_{18}{}^{3+}$.

| **Fragment Ion** | **Charge** | **Theoretical m/z** | **Observed m/z** | **Error [ppm]** |
|---|---|---|---|---|
| $b_{13}{}^{2+}$ | 2 | 731.428531 | 731.428601 | 0.10 |
| $b_{15}{}^{2+}$ | 2 | 838.494403 | 838.499939 | 6.60 |

Table B.10.: IRMPD-fragments assigned to the extraction of the horizontal scan of the CAD-fragment $y_{24}{}^{4+}$.

| **Fragment Ion** | **Charge** | **Theoretical m/z** | **Observed m/z** | **Error [ppm]** |
|---|---|---|---|---|
| GRTLSD | 1 | 630.320567 | 630.319827 | -1.17 |
| $y_{18}{}^{3+}$ | 3 | 699.735763 | 699.735688 | -0.11 |

Table B.11.: IRMPD-fragments assigned to the extraction of the horizontal scan of the CAD-fragment [$y_{59}$-$H_2O$]9+.

| **Fragment Ion** | **Charge** | **Theoretical m/z** | **Observed m/z** | **Error [ppm]** |
|---|---|---|---|---|
| $y_{24}{}^{4+}$ | 4 | 682.381964 | 682.382029 | 0.10 |
| AKIQDKEGIPPDQQRLIFAGKQ-LEDGRTLSDYNIQ-2$H_2O$-$NH_3{}^{5+}$ | 5 | 783.609728 | 783.613100 | 4.30 |

Table B.12.: IRMPD-fragments assigned to the extraction of the horizontal scan of the CAD-fragments [PSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLED/ EVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQ]$^{5+}$.

| **Fragment ion** | **Charge** | **Theoretical m/z** | **Observed m/z** | **Error [ppm]** |
|---|---|---|---|---|

| | | | | |
|---|---|---|---|---|
| QDKEGIPPDQQRLIFA/ PPDQQRLIFAGKQLED$^{3+}$ | 3 | 612.991610 | 612.992183 | 0.93 |
| ? | 1 | | 616.542138 | |
| SDTIENVKAKIQDKEGIP/ PSDTIENVKAKIQDKEGI | 3 | 656.352985 | 656.348575 | -6.72 |
| ENVKAKIQDKEGIP- PDQQRLIFAGKQ-$H_2O$ | 4 | 729.402897 | 729.395403 | -10.27 |
| PSDTIENVKAKIQDKEGIP- PDQQRLIFAGKQLED/ EVEPSDTIENVKAKIQDKE- GIPPDQQRLIFAGKQ-$H_2O$ | 5 | 757.803189 | 757.803150 | -0.05 |
| PSDTIENVKAKIQD- KEGIPPDQQRLIF | 4 | 766.168995 | 766.161562 | -9.70 |
| PSDTIENVKAKIQDK- EGIPPDQQRLIFAG | 4 | 783.922657 | 783.922123 | -0.68 |
| FAGKQLED | 1 | 889.441411 | 889.449165 | 8.72 |
| QDKEGIPPDQQRLIFA/ PPDQQRLIFAGKQLED | 2 | 918.983777 | 918.996204 | 13.52 |
| SDTIENVKAKIQDKEGIP/ PSDTIENVKAKIQDKEGI | 2 | 984.025840 | 984.043141 | 17.58 |

Table B.13.: IRMPD-fragments assigned to the extraction of the horizontal scan of the CAD-fragment $y_{58}{}^{8+}$.

| **Fragment ion** | **Charge** | **Theoretical m/z** | **Observed m/z** | **Error [ppm]** |
|---|---|---|---|---|
| $y_{24}{}^{4+}$ | 4 | 682.381964 | 682.381233 | -1.07 |
| $y_{18}{}^{3+}$ | 3 | 699.735763 | 699.735284 | -0.68 |
| $y_{58}$-$H_2O^{8+}$ | 8 | 814.692066 | 814.703140 | 13.59 |
| $y_{44}{}^{6+}$ | 6 | 832.456416 | 832.460916 | 5.41 |
| $y_{52}{}^{7+}$ | 7 | 841.750534 | 841.756331 | 6.89 |
| $y_{37}$-$H_2O^{5+}$ | 5 | 847.869019 | 847.867768 | -1.48 |
| $y_{37}{}^{5+}$ | 5 | 851.471132 | 851.478360 | 8.49 |
| KEGIPPDQQRLIFAGKQLED- GRTLSDYNIQKESTLHLVL$^{5+}$ | 5 | 887.278288 | 887.275198 | -3.48 |
| $y_{56}{}^{7+}$ | 7 | 907.207860 | 907.198666 | -10.13 |
| $y_{40}{}^{5+}$ | 5 | 913.297627 | 913.306646 | 9.87 |
| ? | 7 | | 918.652379 | |
| ? | 7 | | 921.221214 | |

| | | | | |
|---|---|---|---|---|
| VKAKIQDKEGIPPDQQRLI-FAGKQLEDGRTLSDY-2$H_2O$ | 4 | 951.511430 | 951.519743 | 8.74 |

Table B.14.: IRMPD-fragments assigned to the extraction of the horizontal scan of the CAD-fragment ${b_{16}}^{2+}$.

| **Fragment ion** | **Charge** | **Theoretical m/z** | **Observed m/z** | **Error [ppm]** |
|---|---|---|---|---|
| $b_5$ | 1 | 619.327231 | 619.326172 | -1.71 |
| $b_{16}$-3$H_2O^{2+}$ | 2 | 875.992225 | 875.997994 | 6.59 |
| $b_{16}$-2$H_2O^{2+}$ | 2 | 885.000050 | 885.004288 | 4.79 |
| $b_{16}$-$H_2O^{2+}$ | 2 | 894.007875 | 894.008302 | 0.48 |
| $b_9$ | 1 | 1062.601616 | 1062.610160 | 8.04 |

Table B.15.: CAD fragments assigned to the autocorrelation line of the CAD/ECD MS/2D MS experiment.

| **Precursor Ion** | **Charge** | **Theoretical *m/z*** | **Observed *m/z*** | **Error [ppm]** |
|---|---|---|---|---|
| ? | 1 | | 371.228635 | |
| $b_3$ | 1 | 373.190403 | 373.191501 | 2.94 |
| ${a_8}^{2+}$ | 2 | 467.2831 | 467.282727 | -0.80 |
| ${b_8}^{2+}$ | 2 | 481.280607 | 481.281437 | 1.73 |
| ${y_{18}}^{4+}$ | 4 | 525.053641 | 525.054136 | 0.94 |
| ${y_{24}}^{5+}$ | 5 | 546.107026 | 546.106767 | -0.48 |
| a15$^{3+}$ | 3 | 550.0004 | 550.001098 | 1.27 |
| [$y_{15}$-3$NH_3$]$^{3+}$ | 3 | 552.645774 | 552.644989 | -1.42 |
| [$b_{15}$-$H_2O$]$^{3+}$ | 3 | 553.328505 | 553.328994 | 0.88 |
| PSDTIENVKAKIQDK | 3 | 556.636944 | 556.636753 | -0.34 |
| ${b_{15}}^{3+}$ | 3 | 559.332027 | 559.332576 | 0.98 |
| ? | 2 | | 567.367034 | |
| ${y_{25}}^{5+}$ | 5 | 569.112415 | 569.111575 | -1.48 |
| TLEVE | 1 | 572.292621 | 572.291926 | -1.22 |
| ? | 1 | | 573.296622 | |
| ? | 1 | | 581.24341 | |
| [$b_{16}$-2$H_2O$]$^{3+}$ | 3 | 590.339182 | 590.339491 | 0.52 |
| $a_5$ | 1 | 591.3323 | 591.332947 | 1.09 |
| ${a_{16}}^{3+}$ | 3 | 593.0146 | 593.015322 | 1.22 |

| | | | | |
|---|---|---|---|---|
| $[b_{16}\text{-}H_2O]^{3+}$ | 3 | 596.342703 | 596.343009 | 0.51 |
| ${b_{16}}^{3+}$ | 3 | 602.346225 | 602.345813 | -0.68 |
| $[y_{37}\text{-}H_2O]^{7+}$ | 1 | 605.908521 | 605.90868 | 0.26 |
| ? | 3 | | 607.677227 | |
| ? | 3 | | 607.702009 | |
| ${y_{37}}^{7+}$ | 7 | 608.481459 | 608.482099 | 1.05 |
| PPDQQRLIFAGKQLED/ QDKEGIPPDQQRLIFA | 3 | 612.99161 | 612.992299 | 1.12 |
| IQDKE/KQLED | 1 | 614.314419 | 614.315143 | 1.18 |
| IFAGKQ-28/FAGKQL-28 | 1 | 617.371873 | 617.374276 | 3.89 |
| $b_5$ | 1 | 619.327231 | 619.327877 | 1.04 |
| ? | 3 | | 620.033069 | |
| $[b_{17}\text{-}2H_2O]^{3+}$ | 3 | 623.361986 | 623.362505 | 0.83 |
| ${a_{17}}^{3+}$ | 3 | 626.0374 | 626.037281 | -0.19 |
| $[b_{17}\text{-}H_2O]^{3+}$ | 3 | 629.365508 | 629.366024 | 0.82 |
| ${b_{17}}^{3+}$ | 3 | 635.36903 | 635.370089 | 1.67 |
| ? | 1 | | 637.305845 | |
| $[y_{57}\text{-}NH_3]10+$ | 10 | 642.348233 | 642.352344 | 6.40 |
| $[y_{58}\text{-}H_2O]10+$ | 10 | 651.955108 | 651.95571 | 0.92 |
| ${y_{40}}^{7+}$ | 7 | 652.643241 | 652.641805 | -2.20 |
| ${y_{58}}^{10+}$ | 10 | 653.756164 | 653.756781 | 0.94 |
| ? | 3 | | 656.35274 | |
| ? | 2 | | 661.40001 | |
| $[y_{59}\text{-}H_2O]^{10+}$ | 10 | 664.859367 | 664.860044 | 1.02 |
| $[b_{18}\text{-}2H_2O]^{3+}$ | 3 | 666.376184 | 666.375304 | -1.32 |
| $y_{59}10+$ | 10 | 666.660424 | 666.659099 | -1.99 |
| ${a_{18}}^{3+}$ | 3 | 669.0516 | 669.051987 | 0.58 |
| $y_{53}9+$ | 9 | 669.256765 | 669.250945 | -8.70 |
| QDKEGI/IQDKEG/ GKQLED/KQLEDG | 1 | 671.335883 | 671.334935 | -1.41 |
| $[b_{18}\text{-}H_2O]^{3+}$ | 3 | 672.379706 | 672.3807 | 1.48 |
| ${b_{12}}^{2+}$ | 2 | 674.886499 | 674.886709 | 0.31 |
| ${y_{60}}^{10+}$ | 10 | 676.567265 | 676.569304 | 3.01 |
| ${b_{18}}^{3+}$ | 3 | 678.383228 | 678.383298 | 0.10 |
| ${y_{54}}^{9+}$ | 9 | 681.821661 | 681.819215 | -3.59 |
| ${y_{24}}^{4+}$ | 4 | 682.381964 | 682.381287 | -0.99 |
| $[y_{61}\text{-}H_2O\text{-}NH_3]^{10+}$ | 10 | 685.967813 | 685.965758 | -3.00 |
| KTITLE | 1 | 686.408319 | 686.408773 | 0.66 |

| | | | | |
|---|---|---|---|---|
| $[y_{61}\text{-}NH_3]^{10+}$ | 10 | 687.768869 | 687.769512 | 0.94 |
| ? | 2 | | 689.371421 | |
| $y_{55}{}^{9+}$ | 9 | 693.049181 | 693.049731 | 0.79 |
| $y_{31}{}^{5+}$ | 5 | 694.380171 | 694.379369 | -1.15 |
| ? | 1 | | 696.43827 | |
| TLTGKTI-$H_2O$/LTGKTIT-$H_2O$/ TGKTITL-$H_2O$ | 1 | 697.424303 | 697.42405 | -0.36 |
| ? | 4 | | 699.408149 | |
| $y_{56}{}^{9+}$ | 9 | 705.829952 | 705.829116 | -1.18 |
| $y_{37}{}^{6+}$ | 6 | 709.727156 | 709.727744 | 0.83 |
| $[y_{57}\text{-}NH_3]^{9+}$ | 9 | 713.608339 | 713.612844 | 6.31 |
| ? | 6 | | 715.97408 | |
| ? | 6 | | 716.474126 | |
| $[y_{13}\text{-}H_2O]^{2+}$ | 2 | 716.922793 | 716.923019 | 0.32 |
| $y_{58}\text{-}H_2O^{9+}$ | 9 | 724.282644 | 724.281035 | -2.22 |
| $y_{58}{}^{9+}$ | 9 | 726.283818 | 726.281872 | -2.68 |
| PSDTIENVKAKIQ-DKEGIPPDQQRLI$^{4+}$ | 4 | 729.396275 | 729.397556 | 1.76 |
| $b_{13}{}^{2+}$ | 2 | 731.428531 | 731.429372 | 1.15 |
| $y_{39}\text{-}3NH_3{}^{6+}$ | 6 | 736.5605 | 736.565169 | 6.34 |
| $y_{59}\text{-}2H_2O^{9+}$ | 9 | 736.619537 | 736.619325 | -0.29 |
| $y_{59}\text{-}H_2O^{9+}$ | 9 | 738.62071 | 738.619906 | -1.09 |
| $y_{39}\text{-}2NH_3{}^{6+}$ | 6 | 739.398258 | 739.399219 | 1.30 |
| $?^{2+}$ | 2 | | 741.401234 | |
| GKTITLE | 1 | 743.429783 | 743.429114 | -0.90 |
| $b_6$ | 1 | 747.422194 | 747.422952 | 1.01 |
| $y_{60}\text{-}H_2O\text{-}NH_3{}^{9+}$ | 9 | 747.736473 | 747.737505 | 1.38 |
| $y_{60}\text{-}NH_3{}^{9+}$ | 9 | 749.737647 | 749.737706 | 0.08 |
| $y_{53}\text{-}NH_3{}^{8+}$ | 8 | 750.659633 | 750.661497 | 2.48 |
| $y_{60}{}^{9+}$ | 9 | 751.629486 | 751.627232 | -3.00 |
| $y_{53}{}^{8+}$ | 8 | 752.787951 | 752.777994 | -13.23 |
| $y_{74}\text{-}H_2O^{11+}$ | 11 | 753.962468 | 753.965425 | 3.92 |
| $y_{74}{}^{11+}$ | 11 | 755.599792 | 755.598733 | -1.40 |
| PSDTIENVKAKIQDKEGIP-PDQQRLIFAGKQLED-$H_2O$ EVEPSDTIENVKAKIQDKE-GIPPDQQRLIFAGKQ-$H_2O$ | 5 | 757.803189 | 757.803481 | 0.39 |
| PSDTIENVKAKIQDKEGIPPD$^{3+}$ | 3 | 759.397142 | 759.398046 | 1.19 |

| | | | | |
|---|---|---|---|---|
| PSDTIENVKAKIQDKEGIP-PDQQRLIFAGKQLED EVEPSDTIENVKAKIQDKE-GIPPDQQRLIFAGKQ$^{5+}$ | 5 | 761.405302 | 761.405971 | 0.88 |
| $y_{61}$-$H_2O^{9+}$ | 9 | 763.966378 | 763.968172 | 2.35 |
| $b_{14}$-2$NH_3{}^{2+}$ | 2 | 764.925822 | 764.919332 | -8.48 |
| $?^{2+}$ | 2 | 764.952894 | 764.963273 | |
| $y_7 5^{11+}$ | 11 | 767.241481 | 767.239714 | -2.30 |
| PSDTIENVKAKIQD$^{2+}$ | 2 | 770.404297 | 770.406033 | 2.25 |
| $y_{27}{}^{4+}$ | 2 | 771.670365 | 771.669851 | -0.67 |
| $b_{14}$-$H_2O^{2+}$ | 2 | 772.947088 | 772.945631 | -1.88 |
| $b_{14}$-$NH_3{}^{2+}$ | 2 | 773.439096 | 773.433962 | -6.64 |
| $y_{55}{}^{8+}$ | 8 | 779.554419 | 779.556257 | 2.36 |
| $b_{14}{}^{2+}$ | 2 | 781.952371 | 781.951334 | -1.33 |
| AKIQDKEGIPPDQQRLIFAGKQL-EDGRTLSDYNIQ-2$H_2O$-$NH_3{}^{5+}$ | 5 | 783.609728 | 783.612178 | 3.13 |
| PSDTIENVKAKIQD-KEGIPPDQQRLIFA$^{4+}$ | 4 | 783.922657 | 783.921401 | -1.60 |
| $y_{56}{}^{8+}$ | 8 | 793.932787 | 793.932177 | -0.77 |
| $?^{2+}$ | | | 795.94881 | |
| PSDTIENVKAKIQDK-EGIPPDQQRLIFAG$^{4+}$ | 4 | 798.178023 | 798.177682 | -0.43 |
| $y_{57}$-$H_2O^{8+}$ | 8 | 802.56047 | 802.560793 | 0.40 |
| $y_{28}{}^{4+}$ | 4 | 803.685009 | 803.681829 | -3.96 |
| $y_{57}{}^{8+}$ | 8 | 804.811791 | 804.812648 | 1.07 |
| LEVEPSDTIENVKAKIQDKEGIP-PDQQRLIFAGKQLEDGRTLSDY-NIQKESTLHLVL | 8 | 806.055622 | 806.052677 | -3.65 |
| $y_{43}{}^{6+}$ | 6 | 811.107255 | 811.108253 | 1.23 |
| LTGKTITL-$NH_3$ | 1 | 811.492383 | 811.488497 | -4.79 |
| $y_{58}$-2$H_2O^{8+}$ | 8 | 812.440745 | 812.442111 | 1.68 |
| IFVKTLTGKTITLEV-$H_2O^{2+}$ | 2 | 813.995093 | 813.99604 | 1.16 |
| $y_{58}$-$H_2O^{8+}$ | 8 | 814.692066 | 814.693937 | 2.30 |
| $y_{58}{}^{8+}$ | 8 | 816.943386 | 816.942896 | -0.60 |
| IFVKTLTGKTITLEV$^{2+}$ | 2 | 823.000375 | 823.001739 | 1.66 |
| $y_{59}$-$H_2O^{8+}$ | 8 | 830.82239 | 830.820498 | -2.28 |
| $b_{52}{}^{7+}$ | 7 | 834.452702 | 834.451484 | -1.46 |
| $b_{15}{}^{2+}$ | 2 | 838.494403 | 838.495305 | 1.08 |

| | | | | |
|---|---|---|---|---|
| KTLTGKTI | 1 | 843.529831 | 843.53038 | 0.65 |
| TGKTITLE | 1 | 844.477462 | 844.477061 | -0.47 |
| $y_{37}{}^{5+}$ | 4 | 851.471132 | 851.470268 | -1.01 |
| $b_{31}$-$3H_2O^{4+}$ | 4 | 851.473091 | 851.473419 | 0.39 |
| $y_{15}{}^{2+}$ | 2 | 854.004846 | 854.006549 | 1.99 |
| ? | 1 | | 855.516344 | |
| $b_{39}{}^{5+}$ | 5 | 862.466767 | 862.466937 | 0.20 |
| ? | 1 | | 872.541664 | |
| IFVKTLTGKTITLEVE-28$?^{2+}$ | 2 | 873.521672 | 873.522052 | 0.43 |
| FAGKQLEDGRTLSDYNIQK-ESTLHLVLRLRG-$2H_2O^{4+}$ | 4 | 876.727208 | 876.727099 | -0.12 |
| $?^{2+}$ | 2 | | 879.519458 | |
| $y_{16}$-$3NH_3{}^{2+}$ | 2 | 885.007055 | 885.007815 | 0.86 |
| PSDTIENVKAKIQDKEGIPP-DQQRLIFAGKQLEDGRTLSD$^{5+}$ | 5 | 887.26796 | 887.269708 | 1.97 |
| IFVKTLTGKTITLEVE$^{2+}$ | 2 | 887.521672 | 887.523063 | 1.57 |
| $b_{16}$-$H_2O^{2+}$ | 2 | 894.010417 | 894.008862 | -1.74 |
| ? | 1 | | 899.539645 | |
| $y_{16}$-$NH_3{}^{2+}$ | 2 | 902.033604 | 902.027715 | -6.53 |
| $b_{16}{}^{2+}$ | 2 | 903.0157 | 903.01692 | 1.35 |
| DTIENVKAKIQDKEGI-PPDQQRLIFAGKQLED$^{4+}$ | 4 | 905.48361 | 905.482354 | -1.39 |
| $?^{2+}$ | 2 | | 911.012651 | |
| $y_{40}{}^{5+}$ | 5 | 913.297627 | 913.297331 | -0.32 |
| LEVEPSDTIENVKAKIQDKEGIP-PDQQRLIFAGKQLEDGRTLSDY-NIQKESTLHLVL | 7 | 921.062529 | 921.067695 | 5.61 |
| $b_{17}$-$H_2O^{2+}$ | 2 | 943.544624 | 943.548375 | 3.97 |
| VKAKIQDKEGIPPDQQRLI-FAGKQLEDGRTLSDY-$2H_2O$ | 4 | 951.51143 | 951.505756 | -5.96 |
| $b_{17}{}^{2+}$ | 2 | 952.549907 | 952.550816 | 0.96 |
| $?^{2+}$ | 2 | | 958.529421 | |
| $b_8$ | 1 | 961.553937 | 961.556262 | 2.42 |
| $y_{17}{}^{2+}$ | 2 | 967.568342 | 967.555648 | -13.12 |
| PSDTIENVKAKIQDKEGI/ SDTIENVKAKIQDKEGIP$^{2+}$ | 2 | 984.02584 | 984.027366 | 1.55 |
| $b_{18}$-$2H_2O^{2+}$ | 2 | 999.060639 | 999.063911 | 3.28 |
| $b_{18}$-$H_2O^{2+}$ | 2 | 1008.06592 | 1008.06662 | 0.69 |

| | | | | |
|---|---|---|---|---|
| $b_{18}{}^{2+}$ | 2 | 1017.0712 | 1017.07023 | -0.96 |
| $y_{13}$-2$H_2O$ | 2 | 1414.82775 | 1414.81524 | -8.84 |
| $y_{13}$-$H_2O$ | 2 | 1432.83831 | 1432.84917 | 7.58 |

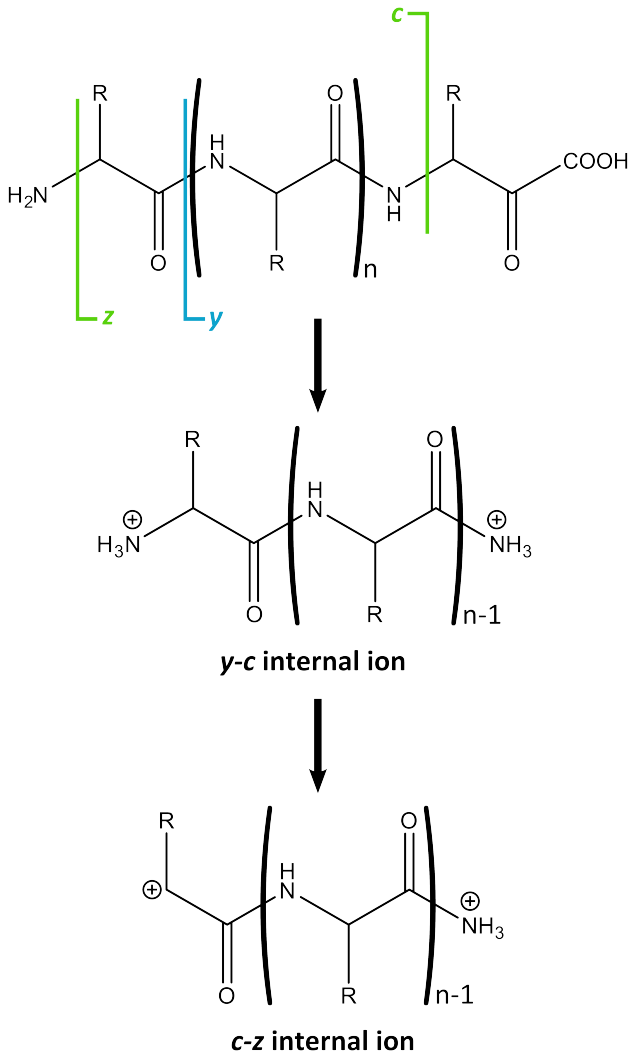



Figure B.1.: Scheme 1: Structure of internal $c$ ions from $y$ ions.

Table B.16.: ECD-fragments assigned to the extraction of the horizontal scan of the CAD-fragment $y_{58}{}^{9+}$. It is hypothesised that the numerous unassigned peak are due to internal fragmentation.

| **Fragment** | **Charge** | **Theoretical m/z** | **Observed m/z** | **Error [ppm]** |
|---|---|---|---|---|
| $?^{3+}$ | 3 | | 556.963948 | |
| $?^{3+}$ | 3 | | 694.254017 | |

| | | | | |
|---|---|---|---|---|
| $?^{3+}$ | 3 | | 699.403801 | |
| $?^{3+}$ | 3 | | 699.733218 | |
| $?^{9+}$ | 9 | | 720.053663 | |
| $[y_{58}\text{-}H_2O]^{9+}$ | 9 | 724.282644 | 724.278964 | |
| $?^{5+}$ | 5 | | 725.926252 | |
| $(z{+}1)_{48}{}^{7+}$ | 7 | 780.714203 | 780.714147 | -0.07 |
| $(z{+}1)_{49}{}^{7+}$ | 7 | 790.862362 | 790.859197 | -4.00 |
| $z_{57}{}^{8+}$ | 8 | 802.809450 | 802.815874 | 8.00 |
| $?^{3+}$ | 3 | | 808.772825 | |
| $(z{+}1)_{50}{}^{7+}$ | 7 | 809.161643 | 809.160035 | -1.99 |
| $c_{57}{}^{8+}$ | 8 | 809.692701 | 809.691508 | -1.47 |
| $c_{50}{}^{7+}$ | 7 | 809.855720 | 809.856491 | 0.95 |
| $y_{58}{}^{8+\bullet}\text{-}CH_3NO$ | 8 | 811.442230 | 811.442373 | 0.18 |
| $y_{58}{}^{8+\bullet}\text{-}CH_3\text{-}H$ | 8 | 815.066207 | 815.066472 | 0.33 |
| $y_{58}{}^{8+\bullet}$ | 8 | 817.069913 | 817.068420 | -1.83 |
| $(z{+}1)_{51}{}^{7+}$ | 7 | 823.314273 | 823.314187 | -0.10 |
| $z_{52}{}^{7+}$ | 7 | 839.462145 | 839.460246 | -2.26 |
| $(z{+}1)_{52}{}^{7+}$ | 7 | 839.606120 | 839.605564 | -0.66 |
| $z_{45}{}^{6+}$ | 6 | 848.957786 | 848.956449 | -1.57 |
| $y_{45}{}^{6+}$ | 6 | 851.627573 | 851.629220 | 1.93 |
| $z_{53}{}^{7+}$ | 7 | 857.896801 | 857.896170 | -0.74 |
| $(z{+}1)_{46}{}^{6+}$ | 6 | 870.468853 | 870.470005 | 1.32 |
| $c_{46}{}^{6+}$ | 6 | 871.626680 | 871.629667 | 3.43 |
| $(z{+}1)_{54}{}^{7+}$ | 7 | 874.195643 | 874.195691 | 0.05 |
| $c_{54}{}^{7+}$ | 7 | 878.618242 | 878.617926 | -0.36 |
| $z_{55}{}^{7+}$ | 7 | 888.487050 | 888.489257 | 2.48 |
| $(z{+}1)_{47}{}^{6+}$ | 6 | 889.316197 | 889.314948 | -1.40 |
| $c_{55}{}^{7+}$ | 7 | 894.773108 | 894.773739 | 0.71 |
| $z_{56}{}^{7+}$ | 7 | 904.919471 | 904.917973 | -1.66 |
| $z_{57}{}^{7+}$ | 7 | 917.352618 | 917.353225 | 0.66 |
| $c_{41}{}^{5+}$ | 5 | 923.285936 | 923.287661 | 1.87 |
| $y_{58}{}^{7+\bullet\bullet}$ | 7 | | 925.216108 | |
| $y_{58}{}^{7+\bullet\bullet}$ | 7 | 927.503123 | 927.505782 | 2.87 |
| $y_{58}{}^{7+\bullet\bullet}\text{-}CH_4N_2$ | 7 | | 931.504462 | |
| $y_{58}{}^{7+\bullet\bullet}$ | 7 | 933.794187 | 933.797634 | 3.69 |
| $c_{53}{}^{6+}$ | 6 | 998.869885 | 998.873582 | 3.70 |
| $(z{+}2)_{56}{}^{6+}$ | 6 | 1055.907445 | 1055.906965 | -0.45 |
| $c_{57}{}^{6+}$ | 6 | 1079.254510 | 1079.245266 | -8.56 |
| ? | 1 | | 1168.687230 | |

Table B.17.: ECD-fragments assigned to the extraction of the horizontal scan of the CAD-fragment $y_{24}{}^{5+}$. It is hypothesised that the numerous unassigned peak are due to internal fragmentation.

| **Fragment** | **Charge** | **Theoretical m/z** | **Observed m/z** | **Error [ppm]** |
|---|---|---|---|---|
| $z_{17}{}^{3+}$ | 3 | 640.041745 | 640.039931 | -2.84 |
| $c_{21}{}^{4+}$ | 4 | 653.625228 | 653.623767 | -2.23 |
| ? | 4 | | 671.375921 | |
| ? | 3 | | 672.376424 | |
| $y_{24}$-$NH_3{}^{4+}$ | 4 | 678.125327 | 678.120719 | -6.80 |
| $?^{3+}$ | 3 | | 759.394042 | |
| $(z{+}1)_{22}{}^{3+}$ | 3 | 833.462369 | 833.459257 | -3.73 |
| ? | 2 | | 894.007495 | |

Table B.18.: ECD-fragments assigned to the extraction of the horizontal scan of the CAD-fragment $y_{24}{}^{4+}$. It is hypothesised that the numerous unassigned peaks are due to internal fragmentation.

| **Fragment** | **Charge** | **Theoretical m/z** | **Observed m/z** | **Error [ppm]** |
|---|---|---|---|---|
| $?^{2+}$ | 2 | | 625.363869 | |
| $?^{2+}$ | 2 | | 838.490214 | |
| $c_{21}{}^{3+}$ | 3 | 871.164545 | 871.160223 | -4.96 |
| $c_{22}{}^{3+}$ | 3 | 890.171700 | 890.173016 | 1.48 |
| $?^{3+}$ | 3 | | 894.832274 | |
| $?^{3+}$ | 3 | | 897.137584 | |
| $z_{24}{}^{3+}$ | 3 | 904.167285 | 904.164248 | -3.36 |
| $c_{23}{}^{3+}$ | 3 | 909.178855 | 909.170201 | -9.52 |
| $?^{2+}$ | 2 | | 952.544795 | |
| $z_{17}{}^{2+}$ | 2 | | 959.555088 | |
| $z_{18}{}^{2+}$ | 2 | | 1041.090123 | |
| y_$c_{71}$ | 2 | 1171.650593 | 1171.643851 | -5.75 |
| $c_{20}2+$ | 2 | 1228.192625 | 1228.187146 | -4.46 |
| $z_{22}{}^{2+}$ | 2 | 1249.186003 | 1249.178612 | -5.92 |
| $c_{13}$ | 1 | 1422.733572 | 1422.720484 | -9.20 |

Table B.19.: ECD-fragments assigned to the extraction of the horizontal scan of the CAD-fragment $y_{58}{}^{8+}$. It is hypothesised that the numerous unassigned peak are due to internal fragmentation.

| Fragment | Charge | Theoretical m/z | Observed m/z | Error [ppm] |
|---|---|---|---|---|
| $y_{24}{}^{5+}$ | 5 | 546.1070264 | 546.1056785 | -2.47 |
| $?^{3+}$ | 3 | | 626.0348669 | |
| $?^{2+}$ | 2 | | 682.3795713 | |
| $?^{2+}$ | 2 | | 694.2544272 | |
| $?^{3+}$ | 3 | | 714.5596568 | |
| $?^{3+}$ | 3 | | 724.3897701 | |
| $(z+1)_{16}{}^{2+}$ | 2 | 903.0414285 | 903.0423467 | 1.02 |
| $(z+1)_{50}{}^{6+}$ | 6 | 943.8540370 | 943.8487402 | -5.61 |
| $?^{6+}$ | 6 | | 944.8279600 | |
| $y_{51}$-$H_2O$-$NH_3{}^{6+}$ | 6 | 957.0277368 | 957.0262922 | -1.51 |
| $(z+1)_{51}{}^{6+}$ | 6 | 960.3654393 | 960.3609651 | -4.66 |
| $y_{17}{}^{2+}$ | 2 | 967.5683420 | 967.5323718 | -37.18 |
| $(z+1)_{43}{}^{5+}$ | 5 | 970.1250714 | 970.1205980 | -4.61 |
| $z_{52}{}^{6+}$ | 6 | 979.2046232 | 979.2010184 | -3.68 |
| $c_{52}{}^{6+}$ | 6 | 980.0225405 | 980.0356960 | 13.42 |
| $(z+1)_{53}{}^{6+}$ | 6 | 1000.8796930 | 1000.8732088 | -6.48 |
| $(z+1)_{45}{}^{5+}$ | 5 | 1018.7494528 | 1018.7393395 | -9.93 |
| $(z+1)_{54}{}^{6+}$ | 6 | 1019.7270370 | 1019.7227455 | -4.21 |
| $?^{6+}$ | 6 | | 1025.0462948 | |
| $(z+1)_{55}{}^{6+}$ | 6 | 1036.5683168 | 1036.5623880 | -5.72 |
| $?^{6+}$ | 6 | | 1043.8978826 | |
| $(z+1)_{46}{}^{5+}$ | 5 | 1044.3611684 | 1044.3548690 | -6.03 |
| $?^{5+}$ | 5 | | 1045.9450104 | |
| $(z+1)_{56}{}^{6+}$ | 6 | 1055.7394742 | 1055.7336105 | -5.55 |
| $c_{47}{}^{5+}$ | 5 | 1063.1569668 | 1063.1521926 | -4.49 |
| ? | 5 | | 1063.3521926 | |
| $(z+1)_{47}{}^{5+}$ | 5 | 1066.9779812 | 1066.9711706 | -6.38 |
| $(z+2)57^{6+}$ | 6 | 1070.4127832 | 1070.4063591 | -6.00 |
| $y_{12}$-$2H_2O$ | 1 | 1285.7851510 | 1285.7966765 | 8.96 |

### B.1.3. MS/MS/2D MS of Ubiquitin

Table B.20.: Fragments assigned to the ISD MS/MS mass spectrum of Ubiquitin in denaturing conditions, the first step of the ISD/CAD/IRMPD (or ECD) MS/MS/2D MS experiment.

| Ion | Charge | Observed m/z | Theoretical m/z | Error [ppm] |
|---|---|---|---|---|
| $b_3$ | 1 | 373.190402 | 373.190403 | 0.00 |
| b4 | 1 | 520.259856 | 520.258817 | 2.00 |
| $b_5$ | 1 | 619.329379 | 619.327231 | 3.47 |
| ${b_{11}}^{2+}$ | 2 | 624.363638 | 624.3626595 | 1.57 |
| ${b_{12}}^{2+}$ | 2 | 674.887856 | 674.886499 | 2.01 |
| ${b_{13}}^{2+}$ | 2 | 731.42925 | 731.428531 | 0.98 |
| ${b_{14}}^{2+}$ | 2 | 781.952438 | 781.9523705 | 0.09 |
| ${b_{15}}^{2+}$ | 2 | 838.493491 | 838.4944025 | -1.09 |
| ${b_{16}}^{2+}$ | 2 | 903.014217 | 903.0156995 | -1.64 |
| ${b_{18}}^{2+}$ | 2 | 1017.0712 | 1017.071204 | 0.00 |
| ${b_{16}}^{3+}$ | 3 | 602.346933 | 602.346225 | 1.18 |
| ${b_{17}}^{3+}$ | 3 | 635.370803 | 635.3690297 | 2.79 |
| ${y_{10}}^{2+}$ | 2 | 567.368536 | 567.3669245 | 2.84 |
| ${y_{11}}^{2+}$ | 2 | 617.89037 | 617.890764 | -0.64 |
| ${y_{12}}^{2+}$ | 2 | 661.407941 | 661.4067785 | 1.76 |
| ${y_{13}}^{2+}$ | 2 | 725.930158 | 725.9280755 | 2.87 |
| ${y_{14}}^{2+}$ | 2 | 789.976556 | 789.975557 | 1.26 |
| ${y_{15}}^{2+}$ | 2 | 854.004811 | 854.004846 | -0.04 |
| ${y_8}^{2+}$ | 2 | 442.295674 | 442.2954365 | 0.54 |
| ${y_9}^{2+}$ | 2 | 510.825415 | 510.8248925 | 1.02 |
| ${y_{14}}^{3+}$ | 3 | 526.987191 | 526.98613 | 2.01 |
| ${y_{15}}^{3+}$ | 3 | 569.672843 | 569.6723227 | 0.91 |
| ${y_{16}}^{3+}$ | 3 | 607.36696 | 607.3670107 | -0.08 |
| ${y_{18}}^{3+}$ | 3 | 699.736956 | 699.735763 | 1.70 |
| ${y_{27}}^{3+}$ | 3 | 1028.5565 | 1028.558061 | -1.52 |
| ${y_{24}}^{4+}$ | 4 | 682.383624 | 682.381964 | 2.43 |
| ${y_{25}}^{4+}$ | 4 | 711.137945 | 711.1387 | -1.06 |
| ${y_{26}}^{4+}$ | 4 | 743.399605 | 743.3993485 | 0.35 |
| ${y_{27}}^{4+}$ | 4 | 771.66909 | 771.6703645 | -1.65 |
| ${y_{40}}^{4+}$ | 4 | 1141.37312 | 1141.370215 | 2.55 |
| ${y_{54}}^{7+}$ | 7 | 876.342514 | 876.3400567 | 2.80 |
| ${y_{39}}^{5+}$ | 5 | 893.888633 | 893.887074 | 1.74 |
| ${y_{40}}^{5+}$ | 5 | 913.294288 | 913.2976268 | -3.66 |
| ${y_{43}}^{5+}$ | 5 | 973.122473 | 973.1272512 | -4.91 |
| ${y_{39}}^{6+}$ | 6 | 745.073023 | 745.0737743 | -1.01 |
| ${y_{40}}^{6+}$ | 6 | 761.249932 | 761.249235 | 0.92 |
| ${y_{49}}^{6+}$ | 6 | 925.00604 | 925.006693 | -0.71 |

| | | | | |
|---|---|---|---|---|
| $y_{50}{}^{6+}$ | 6 | 946.354074 | 946.3558535 | -1.88 |
| $y_{50}{}^{7+}$ | 7 | 811.304776 | 811.3060567 | -1.58 |
| $y_{52}{}^{7+}$ | 7 | 841.7507 | 841.7505341 | 0.20 |
| $y_{53}{}^{7+}$ | 7 | 860.181772 | 860.1851904 | -3.97 |
| $y_{56}{}^{7+}$ | 7 | 907.206423 | 907.20786 | -1.58 |
| $y_{57}{}^{7+}$ | 7 | 919.64018 | 919.641007 | -0.90 |
| $y_{58}{}^{7+}$ | 7 | 933.504411 | 933.5056876 | -1.37 |
| $y_{59}$-$H_2O^{7+}$ | 7 | 951.940085 | 951.9403439 | -0.27 |
| $y_{61}{}^{7+}$ | 7 | 984.525236 | 984.5276307 | -2.43 |
| $y_63^{7+}$ | 7 | 1015.10966 | 1015.11788 | -8.10 |
| $y_{52}{}^{8+}$ | 8 | 736.659293 | 736.6576269 | 2.26 |
| $y_{53}{}^{8+}$ | 8 | 752.785107 | 752.7879511 | -3.78 |
| $y_{54}{}^{8+}$ | 8 | 766.923439 | 766.9234591 | -0.03 |
| $y_{56}{}^{8+}$ | 8 | 793.929099 | 793.932787 | -4.65 |
| $y_{58}{}^{8+}$ | 8 | 816.943363 | 816.9433861 | -0.03 |
| $y_{59}{}^{8+}$ | 8 | 833.074558 | 833.0737104 | 1.02 |
| $y_{58}9+$ | 9 | 726.284038 | 726.2838183 | 0.30 |
| $y_{73}{}^{9+}$ | 9 | 910.720406 | 910.7221214 | -1.88 |
| $y_{74}{}^{9+}$ | 9 | 923.286899 | 923.2870174 | -0.13 |

Table B.21.: Fragments assigned to the autocorrelation line of the MS/MS/2D MS spectrum of Ubiquitin respectively activating the protein with nozzle-skimmer fragmentation (in-source dissociation, ISD), then isolating the ion $y_{58}{}^{9+}$ in the quadrupole and fragmenting it with hexapole-CAD before performing IRMPD 2D MS on the CAD-fragments (ISD/CAD/IRMPD MS/MS/2D MS).

| **Ion** | **Charge** | **Theoretical m/z** | **Observed m/z** | **Error [ppm]** |
|---|---|---|---|---|
| $y_{14}{}^{3+}$ | 3 | 526.986130 | 526.986653 | 0.99 |
| $?^{3+}$ | 3 | | 556.308950 | |
| $?^{3+}$ | 3 | | 563.669028 | |
| $y_{16}{}^{3+}$ | 3 | 607.367011 | 607.367203 | 0.32 |
| QDKEGIPPDQQRLIFA/ PPDQQRLIFAGKQLED | 3 | 612.991610 | 612.991647 | 0.06 |
| GRTLSD/DGRTLS | 1 | 630.320567 | 630.320610 | 0.07 |
| SDTIENVKAKIQDKEGIP/ PSDTIENVKAKIQDKEGI | 3 | 656.352985 | 656.353224 | 0.36 |
| $y_{12}{}^{2+}$ | 2 | 661.406779 | 661.404613 | -3.27 |

| | | | | |
|---|---|---|---|---|
| QDKEGI/IQDKEG/ GKQLED/KQLEDG | 1 | 671.335883 | 671.335863 | -0.03 |
| $y_{24}{}^{4+}$ | 4 | 682.381964 | 682.379788 | -3.19 |
| $y_{37}{}^{6+}$ | 6 | 709.727156 | 709.727302 | 0.21 |
| $y_{25}{}^{4+}$ | 4 | 711.138700 | 711.138485 | -0.30 |
| $y_{44}{}^{7+}$ | 7 | 713.677967 | 713.678409 | 0.62 |
| QQRLIFAGKQLEDGRTLSD/ DQQRLIFAGKQLEDGRTLS/ ENVKAKIQDKEGIPPDQQR-$H_2O$ | 3 | 719.718550 | 719.718324 | -0.31 |
| $y32^{5+}$ | 5 | 723.793853 | 723.794640 | 1.09 |
| $y_{13}{}^{2+}$ | 2 | 725.928076 | 725.928162 | 0.12 |
| $y_{58}9+$ | 9 | 726.283818 | 726.283657 | -0.22 |
| $y_{52}{}^{8+}$ | 8 | 736.657627 | 736.657601 | -0.04 |
| PSDTIENVKAKIQDKEGIPP-DQQRLIFAGKQLEDGRTLSD | 6 | 739.557846 | 739.557370 | -0.64 |
| $y_{26}{}^{4+}$ | 4 | 743.399349 | 743.399229 | -0.16 |
| $y_{40}{}^{6+}$ | 6 | 761.249235 | 761.249269 | 0.04 |
| PSDTIENVKAKIQDKEG-IPPDQQRLIFAGKQLED | 5 | 761.405302 | 761.405526 | 0.29 |
| DQQRLIFAGKQLE/ QQRLIFAGKQLED | 2 | 764.417541 | 764.417586 | 0.06 |
| PSDTIENVKAKIQ-DKEGIPPDQQRLIF | 3 | 766.163378 | 766.163269 | -0.14 |
| $y_{54}{}^{8+}$ | 8 | 766.923459 | 766.923350 | -0.14 |
| $y_{27}{}^{4+}$ | 4 | 771.670365 | 771.671325 | 1.25 |
| $y_{55}{}^{8+}$ | 8 | 779.554419 | 779.554499 | 0.10 |
| PSDTIENVKAKIQD-KEGIPPDQQRLIFA | 4 | 783.922657 | 783.922806 | 0.19 |
| ? | 6 | | 785.909343 | |
| $y_{14}{}^{2+}$ | 2 | 789.975557 | 789.975902 | 0.44 |
| $y_{56}{}^{8+}$ | 8 | 793.932787 | 793.930149 | -3.32 |
| $y_{28}{}^{4+}$ | 4 | 803.685009 | 803.684669 | -0.42 |
| $y_{57}{}^{8+}$ | 8 | 804.811791 | 804.812201 | 0.51 |
| $?^{8+}$ | 8 | | 806.060313 | |
| $y_{44}{}^{6+}$ | 6 | 832.456416 | 832.457205 | 0.95 |
| $y_{15}{}^{2+}$ | 2 | 854.004846 | 854.004552 | -0.34 |
| $y_{55}{}^{7+}$ | 7 | 890.775439 | 890.775324 | -0.13 |
| $y_{39}{}^{5+}$ | 5 | 893.887074 | 893.886928 | -0.16 |

| | | | | |
|---|---|---|---|---|
| GRTLSDYN | 1 | 907.426824 | 907.427078 | 0.28 |
| $y_{40}{}^{5+}$ | 5 | 913.297627 | 913.297788 | 0.18 |
| PPDQQRLIFAGKQLEDGRTLSDYN | 3 | 915.131459 | 915.132135 | 0.74 |
| $?^{2+}$ | 2 | | 918.978604 | |
| $?^{5+}$ | 5 | | 923.478560 | |
| ENVKAKIQD | 1 | 1026.557838 | 1026.557673 | -0.16 |

Table B.22.: IRMPD-fragments assigned to the extraction of the horizontal scan of the fragment QDKEGIPPDQQRLIFA/PPDQQRLIFAGKQLED$^{3+}$ from the ISD/CAD/IRMPD MS/MS/2D MS spectrum.

| **Ion** | **Charge** | **Theoretical m/z** | **Observed m/z** | **Error [ppm]** |
|---|---|---|---|---|
| QDKEGIPPDQQRLIFA/ PPDQQRLIFAGKQLED | 3 | 612.99161 | 612.991474 | -0.22 |
| ? | 2 | | 678.375043 | |

Table B.23.: IRMPD-fragments assigned to the extraction of the horizontal scan of the fragment $y_{24}{}^{4+}$ from the ISD/CAD/IRMPD MS/MS/2D MS spectrum.

| **Ion** | **Charge** | **Theoretical m/z** | **Observed m/z** | **Error [ppm]** |
|---|---|---|---|---|
| GRTLSD-$H_2O$ | 1 | 612.309980 | 612.310002 | -0.04 |
| GRTLSD | 1 | 630.320453 | 630.320567 | -0.18 |
| $y_{12}{}^{2+}$ | 2 | 661.406530 | 661.406779 | -0.38 |
| $[y_{24}$-$H_2O]^{4+}$ | 4 | 677.879478 | | |
| $y_{24}{}^{4+}$ | 4 | 682.382025 | 682.381964 | 0.09 |
| $y_{18}{}^{3+}$ | 3 | 699.735651 | 699.735763 | -0.16 |
| GRTLSDYNIQ | 1 | 1148.566785 | 1148.569466 | -2.33 |

Table B.24.: IRMPD-fragments assigned to the extraction of the horizontal scan of the fragment PSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLED$^{5+}$ from the ISD/CAD/IRMPD MS/MS/2D MS spectrum.

| **Ion** | **Charge** | **Theoretical m/z** | **Observed m/z** | **Error [ppm]** |
|---|---|---|---|---|
| PSDTIENVKAKIQDKEGIPPD | 3 | 759.397142 | 759.396988 | -0.20 |

| | | | | |
|---|---|---|---|---|
| PSDTIENVKAKIQDKEGIP-PDQQRLIFAGKQLED-$H_2O$ | 5 | 757.803189 | 757.802683 | -0.67 |

Table B.25.: IRMPD-fragments assigned to the extraction of the horizontal scan of the fragment ${y_{40}}^{6+}$ from the ISD/CAD/IRMPD MS/MS/2D MS spectrum.

| **Ion** | **Charge** | **Theoretical m/z** | **Observed m/z** | **Error [ppm]** |
|---|---|---|---|---|
| $y_{28}$-$H_2O^{4+}$ | 2 | 799.182368 | 799.180426 | -2.43 |
| ${y_{40}}^{5+}$ | 5 | 913.297627 | 913.295234 | -2.62 |
| $y_{40}$-$H_2O^{6+}$ | 6 | 758.247474 | 758.247466 | -0.01 |

Table B.26.: IRMPD-fragments assigned to the extraction of the horizontal scan of the fragment ${y_{58}}^{9+}$ from the ISD/CAD/IRMPD MS/MS/2D MS spectrum.

| **Ion** | **Charge** | **Theoretical m/z** | **Observed m/z** | **Error [ppm]** |
|---|---|---|---|---|
| GRTLSD/DGRTLS | 1 | 630.320567 | 630.320496 | -0.11 |
| ? | 2 | | 651.397767 | |
| ${y_{12}}^{2+}$ | 2 | 661.406779 | 661.406923 | 0.22 |
| ${y_{24}}^{4+}$ | 4 | 682.381964 | 682.381994 | 0.04 |
| ${y_{18}}^{3+}$ | 3 | 699.735763 | 699.735629 | -0.19 |
| ${y_{13}}^{2+}$ | 2 | 725.928076 | 725.927949 | -0.17 |
| $y_{58}9+$ | 9 | 726.283818 | 726.283973 | 0.21 |
| IENVKAKIQDKEG/ ENVKAKIQDKEGI | 2 | 727.4041 | 727.403884 | -0.30 |
| PSDTIENVKAKIQ-DKEGIPPDQQRLI | 4 | 729.396275 | 729.395768 | -0.69 |
| ? | 2 | | 779.676333 | |
| PPDQQRLIFAGKQLED/ QDKEGIPPDQQRLIFA | 2 | 918.983777 | 918.983769 | -0.01 |
| SDTIENVKAKIQDKEGIP/ PSDTIENVKAKIQDKEGI | 2 | 984.02584 | 984.025841 | 0.00 |

Table B.27.: IRMPD-fragments assigned to the extraction of the horizontal scan of the fragment $y_{15}{}^{2+}$ from the ISD/CAD/IRMPD MS/MS/2D MS spectrum.

| **Ion** | **Charge** | **Theoretical m/z** | **Observed m/z** | **Error [ppm]** |
|---|---|---|---|---|
| ? | 2 | | 794.433687 | |
| ? | 2 | | 822.758908 | |
| $y_{15}$-$H_2O^{2+}$ | 2 | 844.999564 | 844.998982 | -0.69 |
| $y_{15}{}^{2+}$ | 2 | 854.004846 | 854.003945 | -1.05 |

Table B.28.: Fragments assigned to the autocorrelation line of the MS/MS/2D MS spectrum of Ubiquitin respectively activating the protein with nozzle-skimmer fragmentation (in-source dissociation, ISD), then isolating the ion $y_{58}{}^{9+}$ in the quadrupole and fragmenting it with hexapole-CAD before performing ECD 2D MS on the CAD-fragments (ISD/CAD/ECD MS/MS/2D MS).

| **Fragment ion** | **Charge** | **Theoretical *m/z*** | **Observed *m/z*** | **Error [ppm]** |
|---|---|---|---|---|
| $z_4$ | 1 | 386.227220 | 386.230238 | 7.81 |
| $z_5$ | 1 | 542.328331 | 542.329715 | 2.55 |
| $?^{2+}$ | 2 | | 625.367549 | |
| $c_{10}$ | 2 | 647.349127 | 647.348349 | -1.20 |
| $z_{12}{}^{2+}$ | 2 | 653.397417 | 653.398173 | 1.16 |
| $z_6$ | 1 | 655.412395 | 655.413831 | 2.19 |
| $y_{23}{}^{4+}$ | 4 | 668.126598 | 668.114724 | -17.77 |
| $y_{24}{}^{4+}$ | 4 | 682.381964 | 682.383489 | 2.24 |
| $c_{11}$ | 2 | 711.870424 | 711.871024 | 0.84 |
| $z_{13}{}^{2+}$ | 2 | 717.918714 | 717.918835 | 0.17 |
| $c_{12}$ | 2 | 755.386439 | 755.386459 | 0.03 |
| $z_{14}{}^{2+}$ | 2 | 781.966195 | 781.966336 | 0.18 |
| $c_6$ | 1 | 810.410445 | 810.410602 | 0.19 |
| $z_{15}{}^{2+}$ | 2 | 845.995484 | 845.997435 | 2.31 |
| $c_{14}$ | 2 | 862.452310 | 862.451604 | -0.82 |
| $c_{21}$ | 3 | 871.164545 | 871.165224 | 0.78 |
| ? | 3 | | 875.502167 | |
| $z_{23}{}^{3+}$ | 3 | 885.160131 | 885.159635 | -0.56 |
| $c_{22}$ | 3 | 890.171700 | 890.174479 | 3.12 |
| ? | 3 | | 894.835308 | |
| $z_{16}{}^{2+}$ | 2 | 902.537516 | 902.538168 | 0.72 |

| | | | | |
|---|---|---|---|---|
| $c_{15}$ | 2 | 930.981766 | 930.980833 | -1.00 |
| ${z_{17}}^{2+}$ | 2 | 959.558980 | 959.559185 | 0.21 |
| ${y_{17}}^{2+}$ | 2 | 967.568342 | 967.568533 | 0.20 |
| $c_{16}$ | 2 | 987.523798 | 987.525355 | 1.58 |
| $c_{17}$ | 2 | 1037.058005 | 1037.057998 | -0.01 |
| ${z_{18}}^{2+}$ | 2 | 1041.090645 | 1041.090079 | -0.54 |
| $c_{18}$ | 2 | 1093.600037 | 1093.598730 | -1.19 |
| ${z_{19}}^{2+}$ | 2 | 1098.604117 | 1098.603494 | -0.57 |
| ? | 1 | | 1111.644672 | |
| ? | 1 | | 1121.581959 | |
| ${z_{20}}^{2+}$ | 2 | 1142.120131 | 1142.122035 | 1.67 |
| $c_9$ | 1 | 1165.596015 | 1165.594686 | -1.14 |
| $c_{19}$ | 2 | 1171.650593 | 1171.651278 | 0.58 |
| ${z_{21}}^{2+}$ | 2 | 1198.662163 | 1198.663550 | 1.16 |
| $z_{11}$ | 1 | 1218.755528 | 1218.753975 | -1.27 |
| $c_{20}$ | 2 | 1228.192625 | 1228.192366 | -0.21 |
| ${z_{22}}^{2+}$ | 2 | 1249.186003 | 1249.184771 | -0.99 |
| $c_{10}$ | 1 | 1293.690978 | 1293.690974 | 0.00 |
| $z_{12}$ | 1 | 1305.787557 | 1305.787990 | 0.33 |
| $c_{22}$ | 2 | 1334.753912 | 1334.735734 | -13.62 |
| ? | 2 | | 1342.257078 | |
| ? | 2 | | 1347.740523 | |
| ? | 2 | | 1356.252737 | |
| $c_{11}$ | 1 | 1422.733572 | 1422.732586 | -0.69 |
| $c_{12}$ | 1 | 1509.765601 | 1509.761448 | -2.75 |

Table B.29.: ECD-fragments assigned to the extraction of the horizontal scan of the fragment ${y_{24}}^{4+}$ from the ISD/CAD/ECD MS/MS/2D MS spectrum.

| **Fragment ion** | **Charge** | **Theoretical** ***m/z*** | **Observed** ***m/z*** | **Error [ppm]** |
|---|---|---|---|---|
| $z_4$ | 1 | 386.227220 | 386.230238 | 7.81 |
| $z_5$ | 1 | 542.328331 | 542.329715 | 2.55 |
| $?^{2+}$ | 2 | | 625.367549 | |
| $c_{10}$ | 2 | 647.349127 | 647.348349 | -1.20 |
| ${z_{12}}^{2+}$ | 2 | 653.397417 | 653.398173 | 1.16 |
| $z_6$ | 1 | 655.412395 | 655.413831 | 2.19 |
| ${y_{23}}^{4+}$ | 4 | 668.126598 | 668.114724 | -17.77 |
| ${y_{24}}^{4+}$ | 4 | 682.381964 | 682.383489 | 2.24 |

| | | | | |
|---|---|---|---|---|
| $c_{11}$ | 2 | 711.870424 | 711.871024 | 0.84 |
| ${z_{13}}^{2+}$ | 2 | 717.918714 | 717.918835 | 0.17 |
| $c_{12}$ | 2 | 755.386439 | 755.386459 | 0.03 |
| ${z_{14}}^{2+}$ | 2 | 781.966195 | 781.966336 | 0.18 |
| $c_6$ | 1 | 810.410445 | 810.410602 | 0.19 |
| ${z_{15}}^{2+}$ | 2 | 845.995484 | 845.997435 | 2.31 |
| $c_{14}$ | 2 | 862.452310 | 862.451604 | -0.82 |
| $c_{21}$ | 3 | 871.164545 | 871.165224 | 0.78 |
| ? | 3 | | 875.502167 | |
| ${z_{23}}^{3+}$ | 3 | 885.160131 | 885.159635 | -0.56 |
| $c_{22}$ | 3 | 890.171700 | 890.174479 | 3.12 |
| ? | 3 | | 894.835308 | |
| ${z_{16}}^{2+}$ | 2 | 902.537516 | 902.538168 | 0.72 |
| $c_{15}$ | 2 | 930.981766 | 930.980833 | -1.00 |
| ${z_{17}}^{2+}$ | 2 | 959.558980 | 959.559185 | 0.21 |
| ${y_{17}}^{2+}$ | 2 | 967.568342 | 967.568533 | 0.20 |
| $c_{16}$ | 2 | 987.523798 | 987.525355 | 1.58 |
| $c_{17}$ | 2 | 1037.058005 | 1037.057998 | -0.01 |
| ${z_{18}}^{2+}$ | 2 | 1041.090645 | 1041.090079 | -0.54 |
| $c_{18}$ | 2 | 1093.600037 | 1093.598730 | -1.19 |
| ${z_{19}}^{2+}$ | 2 | 1098.604117 | 1098.603494 | -0.57 |
| ? | 1 | | 1111.644672 | |
| ? | 1 | | 1121.581959 | |
| ${z_{20}}^{2+}$ | 2 | 1142.120131 | 1142.122035 | 1.67 |
| $c_9$ | 1 | 1165.596015 | 1165.594686 | -1.14 |
| $c_{19}$ | 2 | 1171.650593 | 1171.651278 | 0.58 |
| ${z_{21}}^{2+}$ | 2 | 1198.662163 | 1198.663550 | 1.16 |
| $z_{11}$ | 1 | 1218.755528 | 1218.753975 | -1.27 |
| $c_{20}$ | 2 | 1228.192625 | 1228.192366 | -0.21 |
| ${z_{22}}^{2+}$ | 2 | 1249.186003 | 1249.184771 | -0.99 |
| $c_{10}$ | 1 | 1293.690978 | 1293.690974 | 0.00 |
| $z_{12}$ | 1 | 1305.787557 | 1305.787990 | 0.33 |
| $c_{22}$ | 2 | 1334.753912 | 1334.735734 | -13.62 |
| ? | 2 | | 1342.257078 | |
| ? | 2 | | 1347.740523 | |
| ? | 2 | | 1356.252737 | |
| $c_{11}$ | 1 | 1422.733572 | 1422.732586 | -0.69 |
| $c_{12}$ | 1 | 1509.765601 | 1509.761448 | -2.75 |

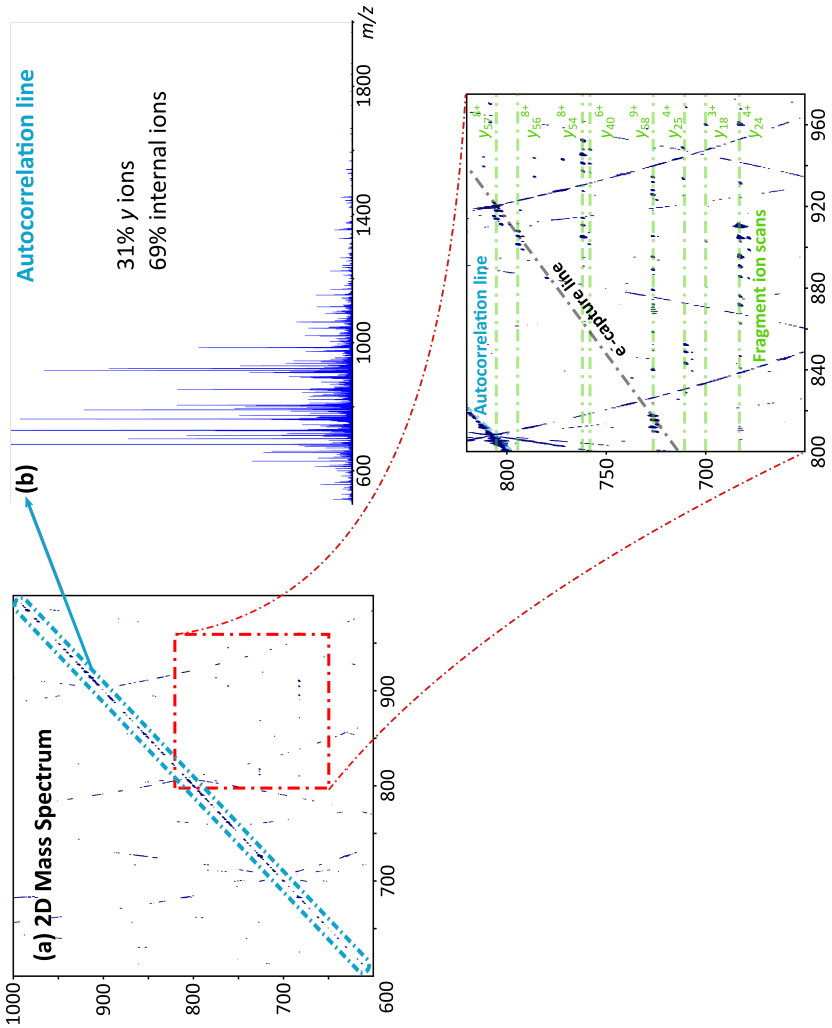



Figure B.2.: (a) 2D mass spectrum of the ISD/CAD/ECD MS/MS/2D MS analysis of Ubiquitin in denaturing conditions. (b) Extraction of the autocorrelation line with percentage of the observed ions. In MS/MS/2D MS, the autocorrelation line gives information equivalent to an MS3 experiment. (c) Zoom of the 2D mass spectrum, with highlight of a fraction of the autocorrelation line, the electron-capture line, and several fragment ion scans. In MS/MS/2D MS, fragment ion scans give information equivalent to MS4 experiments.

Table B.30.: ECD-fragments assigned to the extraction of the horizontal scan of the fragment $y_{58}{}^{9+}$ from the ISD/CAD/ECD MS/MS/2D MS spectrum.

| **Fragment Ion** | **Charge** | **Theoretical m/z** | **Observed m/z** | **Error [ppm]** |
|---|---|---|---|---|
| $y_{18}{}^{3+}$ | 3 | 699.735763 | 699.735525 | -0.3 |
| $y_{13}{}^{2+}$ | 2 | 725.928076 | 725.928574 | 0.7 |
| $y_{58}9+$ | 9 | 726.283818 | 726.284161 | 0.5 |
| $z_{52}{}^{7+}$ | 7 | 839.462145 | 839.461929 | -0.3 |
| c_52 | 7 | 840.163217 | 840.164606 | 1.7 |
| $z_4 5^{6+}$ | 6 | 848.957786 | 848.958566 | 0.9 |
| $z_8$ | 1 | 867.564873 | 867.563841 | -1.2 |
| c_54 | 7 | 878.618242 | 878.618309 | 0.1 |
| $z_{55}{}^{7+}$ | 7 | 888.487050 | 888.485400 | -1.9 |
| c_57 | 7 | 925.219190 | 925.214021 | -5.6 |
| c_53 | 6 | 998.869885 | 998.870031 | 0.1 |
| $z_9$ | 1 | 1004.623785 | 1004.623650 | -0.1 |
| $z_{10}$ | 1 | 1117.707849 | 1117.708271 | 0.4 |
| $z_{11}$ | 1 | 1218.755528 | 1218.756052 | 0.4 |
| $z_{12}$ | 1 | 1305.787557 | 1305.787373 | -0.1 |
| $z_{13}$ | 1 | 1434.830151 | 1434.829955 | -0.1 |
| ? | 7 | | 858.181497 | |

Table B.31.: ECD-fragments assigned to the extraction of the horizontal scan of the fragment $y_{40}{}^{5+}$ from the ISD/CAD/ECD MS/MS/2D MS spectrum.

| **Fragment ion** | **Charge** | **Theoretical m/z** | **Observed m/z** | **Error [ppm]** |
|---|---|---|---|---|
| $c_{16}$-$NH_3$ | 2 | 918.983777 | 918.985819 | 2.22 |
| ? | 4 | | 1116.114945 | |
| $c_{39}$ | 4 | 1126.868845 | 1126.862533 | -5.60 |
| ? | 4 | | 1130.366300 | |
| ? | 4 | | 1137.365032 | |
| CRS4 | 4 | | 1138.869711 | |
| $y_{34}{}^{3+}$ | 3 | 1281.040948 | 1281.041869 | 0.72 |
| $c_{24}$ | 2 | 1323.685362 | 1323.682244 | -2.36 |
| $z_{35}{}^{3+}$ | 3 | 1327.73508 | 1327.73996 | 3.68 |
| $c_{36}$ | 3 | 1393.420488 | 1393.423731 | 2.33 |
| $z_{37}{}^{3+}$ | 3 | 1413.107462 | 1413.106143 | -0.93 |
| $c_{37}$ | 3 | 1431.115176 | 1431.115215 | 0.03 |
| $z_{38}{}^{3+}$ | 3 | 1451.449777 | 1451.453499 | 2.56 |

| | | | | |
|---|---|---|---|---|
| $c_{39}$ | 3 | 1502.156034 | 1502.141239 | -9.85 |
| ? | 3 | | 1507.478830 | |
| CRS3 | 3 | | 1522.492412 | |

### B.1.4. Parameters for MS acquisition of Ubiquitin

Table B.32.: Parameters used for the acquisition of the one-dimensional and two-dimensional mass spectra of Ubiquitin.

| **Parameters** | **FT-ICR MS** | | | | **MS/2D MS** | |
|---|---|---|---|---|---|---|
| | **MS** | **CAD MS/MS** | **CAD/ IRMPD (MS3)** | **CAD/ ECD (MS3)** | **CAD/ IRMPD** | **CAD/ ECD** |
| **Acquisition mode** | | | | | | |
| Acquisition size [datapoints, 16 bit] | 4M | 4M | 4M | 4M | 512k | 512k |
| Horizontal low mass cut-off [$m/z$] | 147.41 | 147.41 | 147.41 | 147.41 | 368.2 | 368.2 |
| Horizontal high mass cut-off [$m/z$] | 3000 | 3000 | 3000 | 3000 | 3000 | 3000 |
| Transient length, $t_2$ [s] | 1.6777 | 1.6777 | 1.6777 | 1.6777 | 0.2097 | 0.2097 |
| Ion accumulation time [s] | 0.2 | 0.2 | 0.2 | 0.2 | 0.2 | 0.2 |
| Ion cooling time [s] | 0.1 | 0.1 | 0.1 | 0.1 | 0.1 | 0.1 |
| **Source (ESI)** | | | | | | |
| Flow rate [uL/h] | 70 | 70 | 70 | 70 | 70 | 70 |
| Capillary voltage [V] | -4800 | -4800 | -4800 | -4800 | -4800 | -4800 |
| EP Off-set [v] | -500 | -500 | -500 | -500 | -500 | -500 |

| | | | | | | |
|---|---|---|---|---|---|---|
| Nebuliser Pressure [bar] | 1.2 | 1.2 | 1.2 | 1.2 | 1.2 | 1.2 |
| Capillary exit [V] | 200 | 200 | 200 | 200 | 200 | 200 |
| Deflector plate [V] | 220 | 220 | 220 | 220 | 220 | 220 |
| Funnel [V] | 150 | 150 | 150 | 150 | 150 | 150 |
| Skimmer [V] | 10 | 10 | 10 | 10 | 10 | 10 |
| Funnel RF [Vpp] | 200 | 200 | 200 | 200 | 200 | 200 |
| **Collision cell** | | | | | | |
| Collision voltage [V] | -4 | -20 | -16 | -16 | -16 | -16 |
| DC extract bias [V] | 0.6 | 0.6 | 0.6 | 0.6 | 0.6 | 0.6 |
| RF frequency [MHz] | 2 | 2 | 2 | 2 | 2 | 2 |
| Collision RF Amplitude [Vpp] | 1400 | 1400 | 1400 | 1400 | 1400 | 1400 |
| **Octopole** | | | | | | |
| Frequency [MHz] | 2 | 2 | 2 | 2 | 2 | 2 |
| RF Amplitude [Vpp] | 1400 | 1400 | 1400 | 1400 | 1400 | 1400 |
| **Transfer optics** | | | | | | |
| Time of Flight [s] | 0.001 | 0.001 | 0.001 | 0.001 | 0.001 | 0.001 |
| Frequency [MHz] | 4 | 4 | 4 | 4 | 4 | 4 |
| RF Amplitude [Vpp] | 350 | 350 | 350 | 350 | 350 | 350 |
| **ICR-cell** | | | | | | |

| | | | | | | |
|---|---|---|---|---|---|---|
| Transfer exit Lens [V] | -15 | -15 | -15 | -15 | -15 | -15 |
| Analyser entrance [V] | -8 | -8 | -8 | -8 | -8 | -8 |
| Side kick [V] | 10 | 10 | 10 | 10 | 10 | 10 |
| Side kick offset [V] | -5 | -5 | -5 | -5 | -5 | -5 |
| Front Trap plate [V] | 0.55 | 0.60 | 0.60 | 0.60 | 0.60 | 0.60 |
| Back Trap plate [V] | 0.57 | 0.62 | 0.62 | 0.62 | 0.62 | 0.62 |
| Sweep Excitation power, $P_3$ [%] | 12 | 12 | 12 | 12 | 12 | 12 |
| **In-cell manipulations** | | | | | | |
| | **multi-CHEF** | | | | **2D MS programme** | |
| $t_1$ [s] | n/a | n/a | n/a | n/a | 1.0E-06 | 1.0E-06 |
| $t_1$ increment [s] | n/a | n/a | n/a | n/a | 1.0E-06 | 1.0E-06 |
| $P_1$ and $P_2$ pulse length [s] | n/a | n/a | n/a | n/a | 4.0E-07 | 4.0E-06 |
| $P_1$ and $P_2$ excitation power [Db] | n/a | n/a | n/a | n/a | 10.42 | 10.42 |
| Loops | n/a | n/a | n/a | n/a | 2048 | 2048 |
| Isolation power [%] | n/a | n/a | 69 | 69 | n/a | n/a |
| Ejection range [$m/z$] | n/a | n/a | 100-3000 | 100-3000 | n/a | n/a |
| Notch width [$m/z$] | n/a | n/a | 4 | 4 | n/a | n/a |
| Shots length [s] | n/a | n/a | 0.158465 | 0.158465 | n/a | n/a |
| Shots power [%] | n/a | n/a | 5 | 5 | n/a | n/a |
| **In-cell Fragmentation** | | | | | | |

| | | | | | | |
|---|---|---|---|---|---|---|
| Laser Power (25W) [%] | / | / | 50 | / | 50 | / |
| IR pulse length [s] | / | / | 0.3 | / | 0.16 | / |
| ECD pulselength [s] | / | / | / | 0.6 | / | 0.016 |
| ECD bias | / | / | / | 1.5 | / | 1.5 |
| ECD Extraction lens | / | / | / | 10 | / | 10 |

## B.2. Mass Spectrometry of Calmodulin

### B.2.1. One-dimensional MS of Calmodulin

Table B.33.: Fragments assigned to the one-dimensional CAD MS/MS spectrum of Calmodulin in denaturing conditions.

| **Ion** | **Charge** | **Observed m/z** | **Theoretical m/z** | **Error [ppm]** |
|---|---|---|---|---|
| y4 | 1 | 450.237927 | 450.238082 | -0.34 |
| AELQD | 1 | 557.256468 | 557.256570 | -0.18 |
| $b_5$ | 1 | 571.272036 | 571.272220 | -0.32 |
| $y_5$ | 1 | 581.278567 | 581.278567 | 0.00 |
| DMINE | 1 | 603.244246 | 603.244291 | -0.07 |
| NPTEAE | 1 | 642.272905 | 642.272949 | -0.07 |
| NEVDAD/ADIDGDG | 1 | 644.252191 | 644.252214 | -0.04 |
| $b_6$-$H_2O$ | 1 | 682.304283 | 682.304249 | 0.05 |
| $y_6$-$NH_3$ | 1 | 692.310630 | 692.310596 | 0.05 |
| ? | 1 | 702.313161 | | |
| $y_6$ | 1 | 709.337144 | 709.337145 | 0.00 |
| $y_{12}$ | 2 | 745.836430 | 745.836347 | 0.11 |
| PTEAELQ | 1 | 769.372692 | 769.372663 | 0.04 |
| $y_{13}$ | 2 | 795.370767 | 795.370554 | 0.27 |
| AELQDMI | 1 | 801.381432 | 801.381119 | 0.39 |
| $y_7$ | 1 | 808.405852 | 808.405559 | 0.36 |
| EAELQDM/DMINEVD | 1 | 817.339856 | 817.339649 | 0.25 |
| $b_7$ | 1 | 829.357023 | 829.357408 | -0.46 |
| ? | 2 | 835.894444 | | |
| LQDMINE/ELQDMIN | 1 | 844.387118 | 844.386933 | 0.22 |
| QNPTEAEL/ NPTEAELQ | 1 | 883.415619 | 883.415591 | 0.03 |
| ? | 1 | 884.400036 | | |
| AELQDMIN | 1 | 915.424053 | 915.424047 | 0.01 |
| LQDMINEV | 1 | 943.455433 | 943.455347 | 0.09 |
| $b_{16}$ | 2 | 946.949345 | 946.949267 | 0.08 |
| $y_8$ | 1 | 955.474257 | 955.473973 | 0.30 |
| $b_{44}$-$NH_3$ | 5 | 968.476188 | 968.472885 | 3.41 |
| ELQDMINE | 1 | 973.429364 | 973.429527 | -0.17 |
| $y_{17}$ | 2 | 973.934603 | 973.934779 | -0.18 |
| $b_{45}$-$H_2O$ | 5 | 994.084640 | 994.084600 | 0.04 |
| $b_{45}$ | 5 | 997.688673 | 997.686713 | 1.96 |

| | | | | |
|---|---|---|---|---|
| NPTEAELQD | 1 | 998.442806 | 998.442535 | 0.27 |
| DMINEVDAD | 1 | 1003.403477 | 1003.403707 | -0.23 |
| $b_{27}$-$H_2O$ | 3 | 1008.477013 | 1008.476739 | 0.27 |
| $b_{46}$ | 5 | 1011.894279 | 1011.894136 | 0.14 |
| ? | 2 | 1013.977790 | | |
| ? | 2 | 1014.976509 | | |
| ? | 1 | 1015.439915 | | |
| $y_{18}$ | 2 | 1031.448051 | 1031.448251 | -0.19 |
| $b_{47}$-$H_2O$ | 5 | 1034.100078 | 1034.100542 | -0.45 |
| $b_{37}$-$H_2O$ | 4 | 1035.758243 | 1035.758406 | -0.16 |
| $b_{47}$ | 5 | 1037.702782 | 1037.702655 | 0.12 |
| ? | 2 | 1039.005216 | | |
| $b_{37}$ | 4 | 1040.261179 | 1040.261048 | 0.13 |
| AELQDMINE/ EAELQDMIN | 1 | 1044.466910 | 1044.466641 | 0.26 |
| b28 | 3 | 1048.163534 | 1048.162820 | 0.68 |
| $b_{48}$-$H_2O$ | 5 | 1056.718425 | 1056.717355 | 1.01 |
| $b_{38}$-$H_2O$ | 4 | 1057.514582 | 1057.516414 | -1.73 |
| $b_{48}$ | 5 | 1060.319520 | 1060.319468 | 0.05 |
| $b_{38}$ | 4 | 1062.019406 | 1062.019055 | 0.33 |
| $b_9$ | 1 | 1070.499793 | 1070.500050 | -0.24 |
| b29-$H_2O$ | 3 | 1075.842010 | 1075.841858 | 0.14 |
| $y_{19}$-$H_2O$ | 2 | 1078.997600 | 1078.985001 | |
| $b_{39}$-$2H_2O$ | 4 | 1081.285086 | 1081.284788 | 0.28 |
| $b_{49}$-$NH_3$ | 5 | 1082.524153 | 1082.525873 | -1.59 |
| $y_9$ | 1 | 1084.516569 | 1084.516567 | 0.00 |
| $b_{39}$-$H_2O$ | 4 | 1085.787276 | 1085.787430 | -0.14 |
| $b_{49}$ | 5 | 1085.931256 | 1085.931183 | 0.07 |
| $y_{19}$ | 2 | 1087.990234 | 1087.990283 | -0.05 |
| $b_{39}$ | 4 | 1090.289860 | 1090.290071 | -0.19 |
| $b_{40}$-$H_2O$ | 4 | 1100.042667 | 1100.042796 | -0.12 |
| $b_{40}$ | 4 | 1104.545595 | 1104.545437 | 0.14 |
| $b_{50}$-$H_2O$ | 5 | 1105.335035 | 1105.334459 | 0.52 |
| $b_{50}$ | 5 | 1108.936466 | 1108.936572 | -0.10 |
| ? | 5 | 1119.945778 | | |
| $b_{19}$ | 2 | 1120.547587 | 1120.541520 | 5.41 |
| ? | 2 | 1121.026046 | | |
| $b_{51}$-$H_2O$ | 5 | 1131.543610 | 1131.542556 | 0.93 |
| $b_{41}$-$H_2O$ | 4 | 1132.057416 | 1132.057440 | -0.02 |
| $b_{51}$ | 5 | 1135.144437 | 1135.144669 | -0.20 |

| | | | | |
|---|---|---|---|---|
| $b_{41}$ | 4 | 1136.559763 | 1136.560081 | -0.28 |
| ? | 4 | 1140.064673 | | |
| $b_{10}$ | 1 | 1141.537034 | 1141.537164 | -0.11 |
| GE**K**LTDEEV ? | 1 | 1143.534992 | | |
| $y_{50}$-2$H_2O$ | 5 | 1148.750259 | 1148.745055 | 4.53 |
| ? | 5 | 1152.162462 | | |
| $b_{52}$-$H_2O$ | 5 | 1154.159419 | 1154.159369 | 0.04 |
| $b_{42}$-2$H_2O$ | 4 | 1156.060163 | 1156.065531 | -4.64 |
| $b_{52}$ | 5 | 1157.761029 | 1157.761482 | -0.39 |
| $b_{42}$-$H_2O$ | 4 | 1160.568790 | 1160.568172 | 0.53 |
| $b_{42}$-$H_2O$ deam? | 4 | 1160.812829 | 1160.814176 | -1.16 |
| $y_{30}$ | 3 | 1164.514272 | 1164.514431 | -0.14 |
| $b_{42}$ | 4 | 1165.070668 | 1165.070813 | -0.12 |
| $b_{31}$ | 3 | 1167.557724 | 1167.557899 | -0.15 |
| ? | 5 | 1168.347567 | | |
| NYEEFVQMM | 1 | 1172.475149 | 1172.475097 | 0.04 |
| ? | 3 | 1173.173708 | | |
| $b_{53}$-$H_2O$ | 5 | 1176.966852 | 1176.967954 | -0.94 |
| $b_{53}$ | 5 | 1180.569687 | 1180.570067 | -0.32 |
| GQNPTEAELQD | 1 | 1183.522644 | 1183.522577 | 0.06 |
| ? | 1 | 1186.545941 | | |
| ELQDMINEVD | 1 | 1187.524749 | 1187.524885 | -0.11 |
| $b_{43}$ | 4 | 1189.333917 | 1189.334004 | -0.07 |
| $b_{54}$-$H_2O$ | 5 | 1202.776165 | 1202.776473 | -0.26 |
| $b_{32}$ | 3 | 1205.252069 | 1205.252587 | -0.43 |
| $b_{54}$ | 5 | 1206.377652 | 1206.378586 | -0.77 |
| ? | 5 | 1209.382380 | | |
| ? | 4 | 1210.344746 | | |
| $y_{10}$ | 1 | 1213.558961 | 1213.559161 | -0.16 |
| $b_{44}$ | 4 | 1214.595986 | 1214.595924 | 0.05 |
| ? | 5 | 1219.181231 | | |
| ? | 5 | 1220.586816 | | |
| $b_{55}$-$H_2O$ | 5 | 1222.587979 | 1222.590156 | -1.78 |
| $b_{55}$ | 5 | 1226.191471 | 1226.192269 | -0.65 |
| $b_{45}$-$H_2O$ | 4 | 1242.353350 | 1242.353931 | -0.47 |
| PTEAELQDMIN/ NPTEAELQDMI | 1 | 1242.566821 | 1242.567084 | -0.21 |
| $b_{56}$-$H_2O$ | 5 | 1245.597006 | 1245.595545 | 1.17 |
| $b_{45}$ | 4 | 1246.855858 | 1246.856573 | -0.57 |
| $b_{56}$ | 5 | 1249.197215 | 1249.197658 | -0.35 |

| | | | | |
|---|---|---|---|---|
| $b_{34}$-$H_2O$ | 3 | 1251.937639 | 1251.938780 | -0.91 |
| ? | 5 | 1256.401097 | | |
| ? | 5 | 1258.005198 | | |
| AELQDMINEVD/ ELQDMINEVDA | 1 | 1258.562157 | 1258.561999 | 0.13 |
| $b_{55}$-$H_2O$ | 5 | 1259.802040 | 1259.802968 | -0.74 |
| $b_{46}$-$H_2O$ | 4 | 1260.112713 | 1260.113210 | -0.39 |
| $b_{55}$ | 5 | 1263.404712 | 1263.405081 | -0.29 |
| $b_{46}$ | 4 | 1264.617946 | 1264.615851 | 1.66 |
| $b_{58}$-$H_2O$ | 5 | 1282.806987 | 1282.808356 | -1.07 |
| $b_{35}$-$H_2O$ | 3 | 1284.961439 | 1284.961584 | -0.11 |
| $b_{58}$ | 5 | 1286.409909 | 1286.410469 | -0.44 |
| $b_{35}$ | 3 | 1290.964525 | 1290.965106 | -0.45 |
| $b_{47}$-$H_2O$ | 4 | 1292.373099 | 1292.373858 | -0.59 |
| $b_{59}$-$NH_3$ | 5 | 1294.413142 | | |
| $b_{47}$ | 4 | 1296.876108 | 1296.876500 | -0.30 |
| $b_{59}$ | 5 | 1297.814850 | 1297.814762 | 0.07 |
| $b_{22}$ | 2 | 1299.616655 | 1299.615946 | 0.55 |
| $y_{91}$ | 8 | 1309.489188 | 1309.490072 | -0.68 |
| $y_{92}$ | 8 | 1318.369091 | 1318.369711 | -0.47 |
| $b_60$ | 5 | 1320.620332 | 1320.623348 | -2.28 |
| $b_{48}$-$H_2O$ | 4 | 1320.645719 | 1320.644874 | 0.64 |
| $b_{48}$ | 4 | 1325.146922 | 1325.147516 | -0.45 |
| $b_61$ | 5 | 1332.026345 | 1332.027641 | -0.97 |
| $b_36$ | 3 | 1334.645760 | 1334.645268 | 0.37 |
| $y_{94}$-$H_2O$ | 8 | 1342.879079 | 1342.880310 | -0.92 |
| $y_{94}$ | 8 | 1345.127751 | 1345.131631 | -2.88 |
| ? | 6 | 1346.299331 | | |
| $b_{49}$-$H_2O$ | 4 | 1352.659724 | 1352.659519 | 0.15 |
| NPTEAELQDMIN | 1 | 1356.610107 | 1356.610012 | 0.07 |
| $b_{49}$ | 4 | 1357.161036 | 1357.162160 | -0.83 |
| $y_{95}$ | 8 | 1361.260171 | 1361.261955 | -1.31 |
| PTEAELQDMINE | 1 | 1371.609391 | 1371.609678 | -0.21 |
| $y_{96}$ | 8 | 1375.515908 | 1375.517321 | -1.03 |
| $b_{24}$ | 2 | 1385.640148 | 1385.640150 | 0.00 |
| $b_{50}$ | 4 | 1385.918432 | 1385.918896 | -0.33 |
| $y_97$-$NH_3$ | 8 | 1387.523731 | 1387.524511 | -0.56 |
| $y_97$ | 8 | 1389.652007 | 1389.652829 | -0.59 |
| MH12 | 12 | 1399.491419 | 1399.328225 | |
| $b_{51}$-$H_2O$ | 4 | 1414.176371 | 1414.176376 | 0.00 |

| | | | | |
|---|---|---|---|---|
| $y_{85}$ | 7 | 1416.810961 | 1416.809822 | 0.80 |
| $b_{51}$ | 4 | 1418.678368 | 1418.679017 | -0.46 |
| $y_{99}$ | 8 | 1420.409916 | 1420.411258 | -0.94 |
| $b_{52}$-$H_2O$ | 4 | 1442.445725 | 1442.447392 | -1.16 |
| AELQDMINEVDAD | 1 | 1444.626205 | 1444.626057 | 0.10 |
| $b_{52}$ | 4 | 1446.948579 | 1446.950033 | -1.01 |
| $y_{88}$ | 7 | 1455.544814 | 1455.545995 | -0.81 |
| $b_{142}$ | 11 | 1462.053385 | 1462.054687 | -0.89 |
| PTEAELQDMINEV | 1 | 1470.676098 | 1470.678092 | -1.36 |
| ? | 5 | 1471.701777 | | |
| $y_{89}$ | 7 | 1471.836657 | 1471.837842 | -0.80 |
| ? | 10 | 1473.204367 | | |
| $b_{53}$ | 4 | 1475.459454 | 1475.460765 | -0.89 |
| $y_{90}$ | 7 | 1479.981875 | 1479.983765 | -1.28 |
| NPTEAELQDMINE | 1 | 1485.652314 | 1485.652606 | -0.20 |
| $y_{12}$ | 1 | 1490.665076 | 1490.665418 | -0.23 |
| $y_{91}$ | 7 | 1496.414728 | 1496.416186 | -0.97 |
| $y_{92}$ | 7 | 1506.561549 | 1506.564345 | -1.86 |
| $b_{54}$ | 4 | 1507.722469 | 1507.721414 | 0.70 |
| ? | 11 | 1511.526555 | | |
| ? | 11 | 1513.163836 | | |
| $y_{93}$ | 7 | 1522.994678 | 1522.996765 | -1.37 |
| $b_{55}$ | 4 | 1532.487252 | 1532.488517 | -0.83 |
| ? | 7 | 1534.431923 | | |
| $y_{94}$ | 7 | 1537.149174 | 1537.149396 | -0.14 |
| ? | 7 | 1539.437860 | | |
| $y_{95}$ | 7 | 1555.583270 | 1555.584052 | -0.50 |
| $b_{56}$ | 4 | 1561.244756 | 1561.245253 | -0.32 |
| ? | 7 | 1569.300076 | | |
| $y_{96}$ | 7 | 1571.874581 | 1571.875899 | -0.84 |
| $b_{55}$ | 4 | 1579.004183 | 1579.004532 | -0.22 |
| NPTEAELQDMINEV | 1 | 1584.720438 | 1584.721020 | -0.37 |
| ? | 7 | 1585.598201 | | |
| $y_{97}$ | 7 | 1588.031568 | 1588.030765 | 0.51 |
| ? | 7 | 1588.141701 | | |
| ? | 7 | 1604.461720 | | |
| ? | 7 | 1604.715737 | | |
| $y_{98}$ | 7 | 1606.747812 | 1606.750835 | -1.88 |
| $b_{59}$-$3H_2O$ | 4 | 1608.008575 | 1608.508710 | |

| | | | | |
|---|---|---|---|---|
| QNPTEAELQDMINE | 1 | 1613.709959 | 1613.711184 | -0.76 |
| $y_{99}$-$NH_3$ ? | 7 | 1620.895391 | 1620.750891 | |
| $b_{59}$ | 4 | 1622.017828 | 1622.016634 | 0.74 |
| $y_{99}$ | 7 | 1623.184649 | 1623.183255 | 0.86 |
| $y_{100}$-$NH_3$ | 7 | 1639.188791 | 1639.044974 | |
| $y_{100}$ | 7 | 1641.475669 | 1641.477338 | -1.02 |
| $b_{60}$ | 4 | 1650.527695 | 1650.527366 | 0.20 |
| PTEAELQDMINEVDA | 1 | 1656.740351 | 1656.742150 | -1.09 |
| GQNPTEAELQDMINE | 1 | 1670.732911 | 1670.732648 | 0.16 |
| ? | 1 | 1682.720795 | | |
| QNPTEAELQDMINEV | 1 | 1712.777743 | 1712.779598 | -1.08 |
| ? | 3 | 1735.165029 | | |
| DGDGQVNYEEFVQMM | 1 | 1743.699284 | 1743.698905 | 0.22 |
| ? | 2 | 1746.266922 | | |
| GQNPTEAELQDMINEV | 1 | 1769.800627 | 1769.801062 | -0.25 |
| PTEAELQDMINEVDAD | 1 | 1771.768377 | 1771.769094 | -0.40 |
| $y_{15}$ | 1 | 1774.812276 | 1774.813874 | -0.90 |
| QNPTEAELQDMINEVD | 1 | 1827.804265 | 1827.806542 | -1.25 |
| ? | 1 | 1867.800147 | | |
| GQNPTEAELQDMINEVD | 1 | 1884.828461 | 1884.828006 | 0.24 |
| NPTEAELQDMINEVDADG/ PTEAELQDMINEVDADGN | 1 | 1942.830820 | 1942.833486 | -1.37 |
| $y_{17}$ | 1 | 1946.862280 | 1946.862282 | 0.00 |
| GQNPTEAELQDMINEVDA | 1 | 1955.863039 | 1955.865120 | -1.06 |
| QNPTEAELQDMINEVDADG/ GQNPTEAELQDMINEVDAD | 1 | 2070.890318 | 2070.892064 | -0.84 |

## B.2.2. MS/2D MS of Calmodulin

### B.2.2.1. CAD/IRMPD MS/2D MS of CaM

The 2D spectrum of figure 4.a was acquired with 2k data-points in the vertical dimension and 512k data-points (16 bit) in the horizontal dimension. The vertical resolving power was RPv $\approx$ 1500, while the horizontal resolving power was $RP_h$ $\approx$ 60000, both reported at $m/z$ 400. The resulting total resolving power for the 2D mass spectrum of Figure 4.a is $RP_{2D}$ $\approx$ 90000000, reported at $m/z$ 400. The instrument was externally calibrated using a standard known mixture. The 2D mass spectrum exhibits a clear autocorrelation line, whose signals correspond with the precursors in correlation with their own cyclotron frequency. The 2D mass spectrum shows also fragment ions from primary CAD ions generated by IRMPD in the ICR-

cell. Harmonic lines and noise lines corresponding to frequencies captured during the acquisition can also be observed. Scintillation noise was minimised using the urQRd algorithm, whose de-noising rank was sufficient for a clear visualisation of the 2D mass spectrum.

Table B.34.: CAD fragments assigned to the autocorrelation line of the CAD/IRMPD MS/2D MS experiment on CaM.

| **Precursor ion** | **Charge** | **Theoretical m/z** | **Observed m/z** | **Error [ppm]** | **Internal** |
|---|---|---|---|---|---|
| y4 | 1 | 450.238082 | 450.237252 | 1.84 | |
| AELQD | 1 | 557.256570 | 557.256391 | 0.32 | 46-50 |
| $b_5$ | 1 | 571.272220 | 571.272281 | -0.11 | |
| $y_5$ | 1 | 581.278567 | 581.277702 | 1.49 | |
| DMINE | 1 | 603.244291 | 603.245342 | -1.74 | 50-54 |
| NPTEAE | 1 | 642.272949 | 642.272972 | -0.04 | 42-47 |
| NEVDAD/ ADIDGDG | 1 | 644.252214 | 644.252256 | -0.07 | |
| $b_6$-$H_2O$ | 1 | 682.304249 | 682.303826 | 0.62 | |
| GQVNYE | 1 | 691.304600 | 691.306375 | -2.57 | 134-139 |
| ? | 2 | | 702.312273 | | |
| $y_6$ | 1 | 709.337145 | 709.326292 | 15.30 | |
| $y_{12}$ | 2 | 745.836347 | 745.836405 | -0.08 | |
| NYEEFV/VNYEEF | 1 | 782.335500 | 782.335278 | 0.28 | |
| $y_{13}$ | 2 | 795.370554 | 795.370698 | -0.18 | |
| AELQDMI | 1 | 801.381119 | 801.381476 | -0.45 | 46-52 |
| $y_7$ | 1 | 808.405559 | 808.404651 | 1.12 | |
| IREADID | 1 | 813.344700 | 813.343908 | 0.97 | 125-131 |
| EAELQDM/ DMINEVD | 1 | 817.339649 | 817.339823 | -0.21 | |
| DGNGTIDF/ GQVNYEE | 1 | 820.347200 | 820.348384 | -1.44 | |
| GQNPTEAE | 1 | 827.353000 | 827.352437 | 0.68 | |
| $b_7$ | 1 | 829.357408 | 829.357339 | 0.08 | |
| QDMINEV | 1 | 830.371300 | 830.372007 | -0.85 | 49-55 |
| ? | 2 | | 835.891340 | | |
| LQDMINE/ ELQDMIN | 1 | 844.386933 | 844.387624 | -0.82 | |
| ? | 1 | | 884.399641 | | |
| LQDMINEV | 1 | 943.455347 | 943.454436 | 0.97 | 48-55 |
| $b_{16}$ | 2 | 946.949267 | 946.950328 | -1.12 | |
| $y_8$ | 1 | 955.473973 | 955.473291 | 0.71 | |
| ? | 6 | | 961.967999 | | |
| $b_{44}$-$H_2O$ | 5 | 968.276100 | 968.275281 | 0.85 | |
| $b_{44}$ | 5 | 971.878200 | 971.876769 | 1.47 | |

| | | | | | |
|---|---|---|---|---|---|
| ELQDMINE | 1 | 973.429527 | 973.430300 | -0.79 | 47-54 |
| DGDGQVNYE | 1 | 978.379900 | 978.379207 | 0.71 | 131-139 |
| $b_{17}$-$H_2O$ | 2 | 981.460000 | 981.463402 | -3.47 | |
| $b_{45}$-$H_2O$ | 5 | 994.084600 | 994.084355 | 0.25 | |
| $b_{45}$ | 5 | 997.686713 | 997.685926 | 0.79 | |
| NPTEAELQD | 1 | 998.442535 | 998.451681 | -9.16 | 42-50 |
| ? | 6 | | 1002.483911 | | |
| ? | 6 | | 1005.483795 | | |
| $b_{27}$-$H_2O$ | 3 | 1008.476739 | 1008.490591 | -13.74 | |
| $b_{46}$ | 5 | 1011.894136 | 1011.894757 | -0.61 | |
| ? | 6 | | 1018.993706 | | |
| ? | 6 | | 1021.995236 | | |
| $b_{47}$-$H_2O$ | 5 | 1034.100542 | 1034.103828 | -3.18 | |
| $b_{47}$ | 5 | 1037.702655 | 1037.702269 | 0.37 | |
| $b_{37}$ | 4 | 1040.261048 | 1040.261333 | -0.27 | |
| AELQDMINE/ EAELQDMIN | 1 | 1044.466641 | 1044.469617 | -2.85 | 46-54/ 45-53 |
| $b_{18}$ | 2 | 1047.007300 | 1047.007948 | -0.62 | |
| ?-$H_2O$ | 6 | | 1050.006841 | | |
| ? | 6 | | 1053.007879 | | |
| $b_{48}$-$H_2O$ | 5 | 1056.717355 | 1056.715775 | 1.49 | |
| $b_{48}$ | 5 | 1060.319468 | 1060.319639 | -0.16 | |
| $b_{38}$ | 4 | 1062.019055 | 1062.021245 | -2.06 | |
| $b_9$ | 1 | 1070.500050 | 1070.507647 | -7.10 | |
| $b_{29}$-$H_2O$ | 3 | 1075.841858 | 1075.843873 | -1.87 | |
| $b_{49}$-$NH_3$ | 5 | 1082.525873 | 1082.527639 | -1.63 | |
| $y_9$ | 1 | 1084.516567 | 1084.513810 | 2.54 | |
| $b_{49}$ | 5 | 1085.931183 | 1085.934533 | -3.08 | |
| $y_{19}$ | 2 | 1087.990283 | 1088.002125 | -10.88 | |
| $b_{39}$ | 4 | 1090.290071 | 1090.290962 | -0.82 | |
| $b_{40}$ | 4 | 1104.545437 | 1104.547310 | -1.70 | |
| $b_{50}$ | 5 | 1108.936572 | 1108.937801 | -1.11 | |
| $b_{19}$ | 2 | 1120.541520 | 1120.541898 | -0.34 | |
| $b_{51}$ | 5 | 1135.144669 | 1135.145339 | -0.59 | |
| $b_{41}$ | 4 | 1136.560081 | 1136.557279 | 2.47 | |
| $b_{10}$ | 1 | 1141.537164 | 1141.538770 | -1.41 | |
| $b_{52}$-$H_2O$ | 5 | 1154.159369 | 1154.158732 | 0.55 | |
| $b_{52}$ | 5 | 1157.761482 | 1157.761959 | -0.41 | |
| $b_{42}$ | 4 | 1165.070813 | 1165.056253 | 12.50 | |

| | | | | | |
|---|---|---|---|---|---|
| $b_{53}$ | 5 | 1180.570067 | 1180.567182 | 2.44 | |
| $b_{54}$ | 5 | 1206.378586 | 1206.371870 | 5.57 | |
| $b_{55}$ | 5 | 1226.192269 | 1226.185247 | 5.73 | |
| $b_{45}$ | 4 | 1246.856573 | 1246.859496 | -2.34 | |
| $b_{47}$ | 4 | 1296.876500 | 1296.874042 | 1.89 | |
| $b_{48}$ | 4 | 1325.147516 | 1325.160317 | -9.66 | |
| $y_{93}$ | 8 | 1332.748100 | 1332.746166 | 1.45 | |
| $b_{51}$ | 4 | 1418.679017 | 1418.676242 | 1.96 | |
| $b_{52}$ | 4 | 1446.950033 | 1446.940278 | 6.74 | |
| NPTEAELQDMINE | 1 | 1485.652606 | 1485.642414 | 6.86 | 42-54 |
| EEEIREAFRVFD | 1 | 1521.7332 | 1521.721712 | 7.55 | 82-93 |
| NPTEAELQDMINEV | 1 | 1584.721020 | 1584.716534 | 2.83 | 42-55 |
| DGDGQVNYEEFVQM | 1 | 1612.658400 | 1612.657381 | 0.63 | 131-144 |
| b29 | 2 | 1622.264400 | 1622.267799 | -2.10 | |
| GDGQVNYEEFVQMM | 1 | 1628.672000 | 1628.678904 | -4.24 | 132-145 |
| PTEAELQDMINEVDA | 1 | 1656.742100 | 1656.731559 | 6.36 | 43-57 |
| GQNPTEAELQDMINE | 1 | 1670.732648 | 1670.731305 | 0.80 | 40-54 |
| NPTEAELQDMINEVD | 1 | 1699.748000 | 1699.743736 | 2.51 | 42-56 |
| QNPTEAELQDMINEV | 1 | 1712.779598 | 1712.803080 | -13.71 | 41-55 |
| DGDGQVNYEEFVQMM-$H_2O$ | 1 | 1725.688300 | 1725.672986 | 8.87 | 131-145 |
| GDGQVNYEEFVQMMT | 1 | 1729.719600 | 1729.719975 | -0.22 | 132-146 |
| DGDGQVNYEEFVQMM | 1 | 1743.698905 | 1743.698595 | 0.18 | 131-145 |
| GQNPTEAELQDMINEV-$NH_3$ NPTEAELQDMINEVDA-$H_2O$ | 1 | 1752.774500 | 1752.783885 | -5.35 | |
| GQNPTEAELQDMINEV | 1 | 1769.801062 | 1769.801708 | -0.36 | 40-55 |
| DGDGQVNYEEFVQMMT | 1 | 1844.746600 | 1844.742435 | 2.26 | 131-146 |
| GQNPTEAELQDMINEVD | 1 | 1884.828006 | 1884.828505 | -0.26 | 40-56 |
| ? | 1 | | 1892.889751 | | |
| TDEEVDEMIREADIDGD-$2H_2O$ | 1 | 1897.775600 | 1897.775904 | -0.16 | 117-133 |
| GQNPTEAELQDMINEVDA | 1 | 1955.865120 | 1955.874241 | -4.66 | 40-57 |
| DIDGDGQVNYEEFVQMM | 1 | 1971.809900 | 1971.799666 | 5.19 | 129-145 |

#### B.2.2.2. CAD/ECD MS/2D MS of CaM

The 2D mass spectrum was acquired with 2k data-points in the vertical dimensions and 1M data-points (16 bit) in the horizontal dimension. The resulting vertical and horizontal resolving powers are respectively RPv $\approx$ 600 and RPh $\approx$ 100000, reported

at $m/z$ 400. The total resolving power was $RP_2D \approx 60$ M, reported at the same $m/z$. Calibration was performed internally during instrumentation setup, using a standard mixture. The 2D mass spectrum of Figure 4.e shows a clear autocorrelation line and peaks corresponding to fragmentation of the precursor ions. Several harmonic lines and electron-capture lines can be observed. Electron-capture lines are highlighted in the spectrum, showing all the precursors that capture one or more electron during the fragmentation period with ECD in the ICR-cell. Scintillation noise was reduced through the urQRd algorithm to a level that made possible to analyse and interpret the 2D mass spectrum.

Table B.35.: CAD fragments assigned to the autocorrelation line of the CAD/ECD MS/2D MS experiment on CaM.

| **Fragment ion** | **Charge** | **Theoretical m/z** | **Observed m/z** | **Error [ppm]** | **Internal** |
|---|---|---|---|---|---|
| PTEAE | 1 | 528.23 | 528.229218 | 1.48 | 43-47 |
| NYEE | 1 | 536.1987 | 536.197937 | 1.42 | 37-40 |
| AELQD | 1 | 557.25657 | 557.256261 | 0.55 | 46-50 |
| $b_5$ | 1 | 571.27222 | 571.271575 | 1.13 | |
| $y_5$ | 1 | 581.278567 | 581.278635 | -0.12 | |
| PTEAEL | 1 | 641.3141 | 641.314357 | -0.40 | 43-48 |
| NPTEAE | 1 | 642.272949 | 642.273193 | -0.38 | 42-47 |
| NEVDAD/ ADIDGDG | 1 | 644.252214 | 644.253124 | -1.41 | |
| $b_6$-$H_2O$ | 1 | 682.304249 | 682.306059 | -2.65 | |
| $y_{11}$ | 2 | 688.8149 | 688.814236 | 0.96 | |
| $y_6$-$NH_3$ | 1 | 692.310596 | 692.312884 | -3.31 | |
| ? | 1 | 702.313161 | 702.313380 | -0.31 | |
| $y_6$ | 1 | 709.337145 | 709.337596 | -0.64 | |
| $y_{12}$ | 2 | 745.836347 | 745.842473 | -8.21 | |
| PTEAELQ | 1 | 769.372663 | 769.376470 | -4.95 | 43-49 |
| NYEEFV/ VNYEEF | 1 | 782.3355 | 782.337189 | -2.16 | |
| $y_{13}$ | 2 | 795.370554 | 795.371245 | -0.87 | |
| AELQDMI | 1 | 801.381119 | 801.384718 | -4.49 | 46-52 |
| $y_7$ | 1 | 808.405559 | 808.410600 | -6.24 | |
| DGNGYISAA-$2H_2O$ ? | 1 | obs 813.36 | 813.349172 | | |
| EAELQDM/DMINEVD | 1 | 817.339649 | 817.341385 | -2.12 | |
| GQNPTEAE | 1 | 827.353 | 827.355943 | -3.56 | 40-47 |
| $b_7$ | 1 | 829.357408 | 829.361369 | -4.78 | |
| ? | 2 | 835.894444 | 835.896468 | -2.42 | |
| LQDMINE/ ELQDMIN | 1 | 844.386933 | 844.391137 | -4.98 | |
| $b_{15}$ | 2 | 873.4151 | 873.416881 | -2.04 | |

| | | | | | |
|---|---|---|---|---|---|
| QNPTEAEL/ NPTEAELQ | 1 | 883.415591 | 883.418610 | -3.42 | |
| ? | 1 | 884.400036 | 884.404264 | -4.78 | |
| QVNYEEF/ NYEEFVQ | 1 | 910.3941 | 910.396928 | -3.11 | |
| AELQDMIN | 1 | 915.424047 | 915.426880 | -3.09 | 46-53 |
| ? | 6 | obs 921.462 | 921.452225 | | |
| $b_{42}$-$H_2O$ | 5 | 928.656000 | 928.660891 | -5.27 | |
| $b_{42}$ | 5 | 932.2581 | 932.261479 | -3.62 | |
| LQDMINEV | 1 | 943.455347 | 943.454829 | 0.55 | 48-55 |
| $b_{16}$ | 2 | 946.949267 | 946.948885 | 0.40 | |
| $y_8$ | 1 | 955.473973 | 955.479018 | -5.28 | |
| ? | 6 | obs 961.984 | 961.970093 | | |
| ? | 5 | obs 962.895 | 962.805820 | | |
| $b_{44}$-$H_2O$ | 5 | 968.2761 | 968.278839 | -2.83 | |
| $b_{44}$ | 5 | 971.8782 | 971.881198 | -3.08 | |
| DGDGQVNYE | 1 | 978.3799 | 978.382787 | -2.95 | 31-39 |
| $b_{17}$-$H_2O$ | 2 | 981.46 | 981.458885 | 1.14 | |
| $b_{45}$-$2H_2O$ | 5 | 990.4825 | 990.481702 | 0.81 | |
| $b_{45}$-$H_2O$ | 5 | 994.0846 | 994.088114 | -3.53 | |
| $b_{45}$ | 5 | 997.686713 | 997.690910 | -4.21 | |
| NPTEAELQD | 1 | 998.442535 | 998.445114 | -2.58 | 42-50 |
| ? | 6 | obs 1002.5 | 1002.486214 | | |
| $b_{27}$-$H_2O$ | 3 | 1008.47674 | 1008.477937 | -1.19 | |
| $b_{46}$ | 5 | 1011.89414 | 1011.897645 | -3.47 | |
| $y_{18}$ | 2 | 1031.44825 | 1031.446932 | 1.28 | |
| $b_{47}$-$H_2O$ | 5 | 1034.10054 | 1034.105283 | -4.58 | |
| $b_{37}$-$H_2O$ | 4 | 1035.75841 | 1035.764857 | -6.23 | |
| $b_{47}$ | 5 | 1037.70265 | 1037.705807 | -3.04 | |
| $b_{37}$ | 4 | 1040.26105 | 1040.265825 | -4.59 | |
| AELQDMINE/ EAELQDMIN | 1 | 1044.46664 | 1044.471924 | -5.06 | |
| ? | 6 | obs 1050.03 | 1050.012127 | | |
| ? | 6 | obs 1056.03 | 1055.475919 | | |
| $b_{48}$-$H_2O$ | 5 | 1056.71735 | 1056.720151 | -2.65 | |
| $b_{38}$-$H_2O$ | 4 | 1057.51641 | 1057.522508 | -5.76 | |
| $b_{48}$ | 5 | 1060.31947 | 1060.322963 | -3.30 | |
| $b_{38}$ | 4 | 1062.01905 | 1062.021917 | -2.70 | |
| $y_9$-$H_2O$ | 1 | 1066.5 | 1066.504788 | -4.49 | |
| $b_9$ | 1 | 1070.50005 | 1070.501031 | -0.92 | |
| ? | 2 | 1076.19 | 1076.178094 | 11.06 | |
| $b_{39}$-$2H_2O$ | 4 | 1081.28479 | 1081.288302 | -3.25 | |

| | | | | | |
|---|---|---|---|---|---|
| $b_{49}$-$H_2O$ | 5 | 1082.32907 | 1082.326989 | 1.92 | |
| $y_9$ | 1 | 1084.51657 | 1084.520372 | -3.51 | |
| $b_{39}$-$H_2O$ | 4 | 1085.78743 | 1085.791275 | -3.54 | |
| $b_{49}$ | 5 | 1085.93118 | 1085.933687 | -2.31 | |
| $b_{39}$ | 4 | 1090.29007 | 1090.294277 | -3.86 | |
| ? | 4 | obs 1093.57 | 1093.798130 | | |
| $b_{40}$-$H_2O$ | 4 | 1100.0428 | 1100.048668 | -5.34 | |
| $b_{40}$ | 4 | 1104.54544 | 1104.549989 | -4.12 | |
| $b_{50}$ | 5 | 1108.93657 | 1108.941953 | -4.85 | |
| $b_{19}$ | 2 | 1120.54152 | 1120.542982 | -1.30 | |
| $b_{51}$-$H_2O$ | 5 | 1131.54256 | 1131.546635 | -3.60 | |
| $b_{41}$-$H_2O$ | 4 | 1132.05744 | 1132.061280 | -3.39 | |
| $b_{51}$ | 5 | 1135.14467 | 1135.150301 | -4.96 | |
| $b_{41}$ | 4 | 1136.56008 | 1136.564219 | -3.64 | |
| $b_{10}$ | 1 | 1141.53716 | 1141.530251 | 6.06 | |
| $y_{90}$-$H_2O$ | 9 | 1149.3211 | 1149.332209 | -9.67 | |
| $b_{52}$-$H_2O$ | 5 | 1154.15937 | 1154.164167 | -4.16 | |
| $b_{52}$ | 5 | 1157.76148 | 1157.766645 | -4.46 | |
| $b_{42}$-$H_2O$ | 4 | 1160.56817 | 1160.575753 | -6.53 | |
| $b_{42}$-$H_2O$ deamidated | 4 | 1160.81418 | 1160.821279 | -6.12 | |
| $y_{90}$ | 9 | 1164.1031 | 1164.107863 | -4.09 | |
| $b_{42}$ | 4 | 1165.07081 | 1165.075349 | -3.89 | |
| $y_{92}$-$NH_3$ | 9 | 1170.1043 | 1170.111609 | -6.25 | |
| $y_{92}$ | 9 | 1171.9961 | 1172.001147 | -4.31 | |
| $b_{53}$ | 5 | 1180.57007 | 1180.580754 | -9.05 | |
| $y_{93}$-$H_2O$ | 9 | 1182.7757 | 1182.778857 | -2.67 | |
| $y_{93}$ | 9 | 1184.7769 | 1184.780062 | -2.67 | |
| $y_{96}$-$2NH_3$ | 9 | 1219.0103 | 1219.016379 | -4.99 | |
| $y_{96}$-$NH_3$ | 9 | 1220.9021 | 1220.911673 | -7.84 | |
| $b_{55}$-$H_2O$ | 5 | 1222.59016 | 1222.574650 | 12.68 | |
| $b_{55}$ | 5 | 1226.19227 | 1226.200988 | -7.11 | |
| PTEAELQDMIN/ NPTEAELQDMI | 1 | 1242.56708 | 1242.554454 | 10.16 | |
| $b_{45}$ | 4 | 1246.85657 | 1246.861835 | -4.22 | |
| ? | 2 | obs 1247.38 | 1247.279620 | | |
| $b_{55}$ | 5 | 1263.40508 | 1263.383515 | 17.07 | |
| $b_{46}$ | 4 | 1264.61585 | 1264.621188 | -4.22 | |
| $b_{47}$ | 4 | 1296.8765 | 1296.878982 | -1.91 | |
| $y_{91}$-$2NH_3$ | 8 | 1305.2334 | 1305.246159 | -9.77 | |

| | | | | | |
|---|---|---|---|---|---|
| $y_{91}$-$NH_3$ | 8 | 1307.3617 | 1307.374492 | -9.78 | |
| $y_{91}$ | 8 | 1309.49007 | 1309.503065 | -9.92 | |
| $y_{92}$ | 8 | 1318.36971 | 1318.371954 | -1.70 | |
| $b_{48}$ | 4 | 1325.14752 | 1325.118833 | 21.64 | |
| $y_{93}$-$NH_3$ | 8 | 1330.6198 | 1330.628293 | -6.38 | |
| $y_{93}$ | 8 | 1332.7481 | 1332.753610 | -4.13 | |
| $y_{94}$-$H_2O$ | 8 | 1342.88031 | 1342.886148 | -4.35 | |
| $y_{96}$ | 8 | 1375.51732 | 1375.517607 | -0.21 | |
| $b_{50}$ | 4 | 1385.9189 | 1385.928876 | -7.20 | |
| $b_{51}$ | 4 | 1418.67902 | 1418.684627 | -3.95 | |
| ? | 7 | obs 1440.15 | 1439.973450 | | |
| $b_{52}$ | 4 | 1446.95003 | 1446.963298 | -9.17 | |
| ? | 11 | 1462.05339 | 1462.052908 | 0.33 | |
| PTEAELQDMINEV | 1 | 1470.67809 | 1470.678523 | -0.29 | 43-55 |
| $y_{89}$ | 7 | 1471.83784 | 1471.834649 | 2.17 | |
| $y_{90}$-$3NH_3$ | 7 | 1475.119 | 1475.122049 | -2.07 | |
| $y_{90}$-$NH_3$ | 7 | 1477.5514 | 1477.554798 | -2.30 | |
| $y_{90}$ | 7 | 1479.98377 | 1479.980926 | 1.92 | |
| FDKDGNGYISAAEL | 1 | 1481.6907 | 1481.702091 | -7.69 | |
| NPTEAELQDMINE | 1 | 1485.65261 | 1485.658412 | -3.91 | 42-54 |
| $y_{12}$ | 1 | 1490.66542 | 1490.671395 | -4.01 | |
| $b_{54}$ | 4 | 1507.72141 | 1507.723659 | -1.49 | |
| ? | 6 | 1528.61 | 1528.442091 | | |
| $y_{94}$ | 7 | 1537.1494 | 1537.147680 | 1.12 | |
| NPTEAELQDMINEV | 1 | 1584.72102 | 1584.718120 | 1.83 | 42-55 |
| $b_{59}$-$3H_2O$ | 4 | 1608.50871 | 1608.514836 | -3.81 | |
| ? | 1 | | 1612.727173 | | |
| GQNPTEAELQDMINE | 1 | 1670.73265 | 1670.739351 | -4.01 | 40-54 |
| NPTEAELQDMINEVD | 1 | 1699.74374 | 1699.752237 | -5.00 | 42-56 |
| GDGQVNYEEFVQMMT | 1 | 1729.71998 | 1729.726696 | -3.89 | 132-146 |
| DGDGQVNYEEFVQMM | 1 | 1743.69891 | 1743.706542 | -4.38 | 131-145 |
| $y_{30}$ | 2 | 1746.268009 | 1746.27271 | -2.69 | |
| GQNPTEAELQDMINEV | 1 | 1769.80106 | 1769.801488 | -0.24 | 40-55 |
| PTEAELQDMINEVDAD | 1 | 1771.76909 | 1771.765980 | 1.76 | 43-58 |
| GQNPTEAELQDMINEVD | 1 | 1884.82801 | 1884.842019 | -7.43 | 40-56 |
| AELQDMINEVDADGNGTI | 1 | 1886.8436 | 1886.825341 | 9.68 | 46-63 |
| QNPTEAELQDMINEVDA | 1 | 1898.8436 | 1898.838924 | 2.46 | 41-57 |
| DMINEVDADGNGTIDFPE | 1 | 1933.812 | 1933.812290 | -0.15 | 50-67 |
| $y_{17}$ | 1 | 1946.86228 | 1946.864204 | -0.99 | |

| | | | | | |
|---|---|---|---|---|---|
| GQNPTEAELQDMINEVDA | 1 | 1956.86392 | 1956.880434 | -8.44 | 40-57 |
| DIDGDGQVNYEEFVQMM | 1 | 1971.79967 | 1971.816153 | -8.36 | 129-145 |

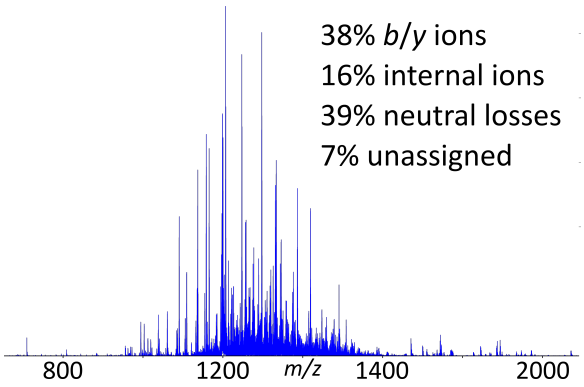



Figure B.3.: Autocorrelation line extracted from the CAD/ECD MS/2D MS spectrum of CaM.

### B.2.3. Parameters for MS acquisition of Calmodulin

Table B.36.: Parameters used for the acquisition of the one-dimensional and two-dimensional mass spectra of Calmodulin.

| **Parameters** | **FT-ICR MS** | | **MS/2D MS** | |
|---|---|---|---|---|
| | **MS** | **CAD MS/MS** | **CAD/IRMPD** | **CAD/ECD** |
| **Acquisition mode** | | | | |
| Acquisition size [datapoints, 16 bit] | 4M | 4M | 512k | 512k |
| Horizontal low mass cut-off [$m/z$] | 147.41 | 147.41 | 368.2 | 368.2 |
| Horizontal high mass cut-off [$m/z$] | 3000 | 3000 | 3000 | 3000 |
| Transient length, $t_2$ [s] | 1.6777 | 1.6777 | 0.2097 | 0.2097 |
| Ion accumulation time [s] | 0.2 | 0.2 | 0.2 | 0.2 |

| | | | | |
|---|---|---|---|---|
| Ion cooling time [s] | 0.1 | 0.1 | 0.1 | 0.1 |
| **Source (ESI)** | | | | |
| Flow rate [uL/h] | 70 | 70 | 70 | 70 |
| Capillary voltage [V] | -4800 | -4800 | -4800 | -4800 |
| EP Off-set [v] | -500 | -500 | -500 | -500 |
| Nebuliser Pressure [bar] | 1.2 | 1.2 | 1.2 | 1.2 |
| Capillary exit [V] | 200 | 200 | 200 | 200 |
| Deflector plate [V] | 220 | 220 | 220 | 220 |
| Funnel [V] | 150 | 150 | 150 | 150 |
| Skimmer [V] | 10 | 10 | 10 | 10 |
| Funnel RF [Vpp] | 200 | 200 | 200 | 200 |
| **Collision cell** | | | | |
| Collision voltage [V] | -4 | -13.5 | -13.5 | -13.5 |
| DC extract bias [V] | 0.6 | 0.6 | 0.6 | 0.6 |
| RF frequency [MHz] | 2 | 2 | 2 | 2 |
| Collision RF Amplitude [Vpp] | 1400 | 1400 | 1400 | 1400 |
| **Octopole** | | | | |
| Frequency [MHz] | 2 | 2 | 2 | 2 |
| RF Amplitude [Vpp] | 1400 | 1400 | 1400 | 1400 |
| **Transfer optics** | | | | |
| Time of Flight [s] | 0.001 | 0.001 | 0.0012 | 0.001 |
| Frequency [MHz] | 4 | 4 | 4 | 4 |

| | | | | |
|---|---|---|---|---|
| RF Amplitude [Vpp] | 350 | 350 | 350 | 350 |
| **ICR-cell** | | | | |
| Transfer exit Lens [V] | -15 | -15 | -15 | -15 |
| Analyser entrance [V] | -8 | -8 | -8 | -8 |
| Side kick [V] | 5 | 12 | 10 | 5 |
| Side kick offset [V] | -1.5 | -2.3 | -2.9 | -1.5 |
| Front Trap plate [V] | 0.51 | 0.51 | 0.69 | 0.91 |
| Back Trap plate [V] | 0.55 | 0.55 | 0.74 | 0.96 |
| Sweep Excitation power, $P_3$ [%] | 16 | 16 | 16 | 16 |
| **In-cell manipulations** | | | | |
| | **multi-CHEF** | | **2D MS pulse programme** | |
| $t_1$ [s] | n/a | n/a | 1.00E-06 | 1.00E-06 |
| $t_1$ increment [s] | n/a | n/a | 1.00E-06 | 1.00E-06 |
| $P_1$ and $P_2$ pulse length [s] | n/a | n/a | 4.00E-07 | 4.00E-06 |
| $P_1$ and $P_2$ excitation power [Db] | n/a | n/a | 10.42 | 10.42 |
| Loops | n/a | n/a | 2048 | 2048 |
| **In-cell Fragmentation** | | | | |
| Laser Power (25W) [%] | / | / | 50 | / |
| IR pulse length [s] | / | / | 0.16 | / |
| ECD pulselength [s] | / | / | / | 0.016 |
| ECD bias | / | / | / | 1.5 |

| | | | | |
|---|---|---|---|---|
| ECD Extraction lens | / | / | / | 10 |

# C. Supporting Information for Chapter 5

## C.1. Principles of 2D MS

## C.2. One-Dimesional control spectra

One-dimensional MS spectra were acquired as controls for each sample, generating full (MS) mass spectra of the polymer mixtures. The control spectra were acquired using 4M data-points (16 bits) for a transient length of 1.6777 s, over a mass range of m/z 147.5-3000. MS control spectra were accumulated for 64 scans.

## C.3. 2D MS of TPGS

Table C.1.: Peaks assigned to the autocorrelation line of the IRMPD 2D MS experiment of TPGS.

| ***n* units of PEG** | **Charge** | **Theorethical m/z** | **Observed m/z** | **Error [ppm]** |
|---|---|---|---|---|
| **Free PEG** | | | | |
| 13 | 1 | 613.340580 | 613.340464 | -0.19 |
| 14 | 1 | 657.366795 | 657.366263 | -0.81 |
| 15 | 1 | 701.393010 | 701.392958 | -0.07 |
| 16 | 1 | 745.419225 | 745.419210 | -0.02 |
| 17 | 1 | 789.445440 | 789.445129 | -0.39 |
| 18 | 1 | 833.471655 | 833.470705 | -1.14 |
| 19 | 1 | 877.497870 | 877.498432 | 0.64 |
| 20 | 1 | 921.524085 | 921.525436 | 1.47 |
| 21 | 1 | 965.550300 | 965.550910 | 0.63 |
| 22 | 1 | 1009.576515 | 1009.575653 | -0.85 |
| 23 | 1 | 1053.602730 | 1053.602237 | -0.47 |
| 24 | 1 | 1097.628945 | 1097.631542 | 2.37 |
| 25 | 1 | 1141.655160 | 1141.657449 | 2.01 |
| 26 | 1 | 1185.681375 | 1185.681979 | 0.51 |
| 27 | 1 | 1229.707590 | 1229.697956 | -7.83 |
| 25 | 2 | 582.322190 | 582.322064 | -0.22 |
| 26 | 2 | 604.335298 | 604.334797 | -0.83 |
| 28 | 2 | 648.361513 | 648.361180 | -0.51 |
| 29 | 2 | 670.374620 | 670.376568 | 2.91 |
| **TPGS** | | | | |
| 25 | 3 | 566.673387 | 566.673261 | -0.22 |
| 13 | 2 | 574.358180 | 574.358334 | 0.27 |
| 26 | 3 | 581.348792 | 581.348782 | -0.02 |

| | | | | |
|---|---|---|---|---|
| 27 | 3 | 596.024197 | 596.023865 | -0.56 |
| 14 | 2 | 596.371288 | 596.360548 | -18.01 |
| 28 | 3 | 610.699602 | 610.699312 | -0.47 |
| 15 | 2 | 618.384395 | 618.384579 | 0.30 |
| 29 | 3 | 625.375007 | 625.374997 | -0.02 |
| 30 | 3 | 640.050412 | 640.049712 | -1.09 |
| 16 | 2 | 640.397503 | 640.398355 | 1.33 |
| 31 | 3 | 654.725817 | 654.725679 | -0.21 |
| 17 | 2 | 662.410610 | 662.409984 | -0.94 |
| 32 | 3 | 669.401222 | 669.401337 | 0.17 |
| 33 | 3 | 684.076627 | 684.076950 | 0.47 |
| 18 | 2 | 684.423718 | 684.423994 | 0.40 |
| 34 | 3 | 698.752032 | 698.751114 | -1.31 |
| 19 | 2 | 706.436825 | 706.437285 | 0.65 |
| 20 | 2 | 728.449933 | 728.450343 | 0.56 |
| 21 | 2 | 750.463040 | 750.463304 | 0.35 |
| 22 | 2 | 772.476148 | 772.476865 | 0.93 |
| 23 | 2 | 794.489255 | 794.489837 | 0.73 |
| 24 | 2 | 816.502363 | 816.501778 | -0.72 |
| 25 | 2 | 838.515470 | 838.515029 | -0.53 |
| 26 | 2 | 860.528578 | 860.528518 | -0.07 |
| 27 | 2 | 882.541685 | 882.541340 | -0.39 |
| 28 | 2 | 904.554793 | 904.553991 | -0.89 |
| 29 | 2 | 926.567900 | 926.567500 | -0.43 |
| 30 | 2 | 948.581008 | 948.581984 | 1.03 |
| 12 | 1 | 1081.700925 | 1081.702793 | 1.73 |
| 13 | 1 | 1125.727140 | 1125.726561 | -0.51 |
| 14 | 1 | 1169.753355 | 1169.754042 | 0.59 |
| 15 | 1 | 1213.779570 | 1213.782210 | 2.17 |
| 16 | 1 | 1257.805785 | 1257.804338 | -1.15 |
| 17 | 1 | 1301.832000 | 1301.830920 | -0.83 |
| 18 | 1 | 1345.858215 | 1345.857404 | -0.60 |
| 19 | 1 | 1389.884430 | 1389.882872 | -1.12 |
| 20 | 1 | 1433.910645 | 1433.912440 | 1.25 |
| 21 | 1 | 1477.936860 | 1477.934847 | -1.36 |
| 22 | 1 | 1521.963075 | 1521.963139 | 0.04 |
| 23 | 1 | 1565.989290 | 1565.992966 | 2.35 |
| 24 | 1 | 1610.015505 | 1610.014519 | -0.61 |
| 25 | 1 | 1654.041720 | 1654.042703 | 0.59 |

| | | | | |
|---|---|---|---|---|
| 26 | 1 | 1698.067935 | 1698.065391 | -1.50 |
| 27 | 1 | 1742.094150 | 1742.098239 | 2.35 |
| 28 | 1 | 1786.120365 | 1786.118034 | -1.30 |
| 29 | 1 | 1830.146580 | 1830.145445 | -0.62 |
| 30 | 1 | 1874.172795 | 1874.171402 | -0.74 |
| 31 | 1 | 1918.199010 | 1918.197169 | -0.96 |
| 32 | 1 | 1962.225225 | 1962.228629 | 1.73 |
| 33 | 1 | 2006.251440 | 2006.252997 | 0.78 |
| **Di-TPGS** | | | | |
| 17 | 2 | 940.616998 | 940.616967 | -0.03 |
| 18 | 2 | 962.630105 | 962.630127 | 0.02 |
| 19 | 2 | 984.643213 | 984.643588 | 0.38 |
| 20 | 2 | 1006.656320 | 1006.655580 | -0.74 |
| 21 | 2 | 1028.669428 | 1028.668736 | -0.67 |
| 22 | 2 | 1050.682535 | 1050.681555 | -0.93 |
| 23 | 2 | 1072.695643 | 1072.696550 | 0.85 |
| 24 | 2 | 1094.708750 | 1094.710564 | 1.66 |
| 25 | 2 | 1116.721858 | 1116.720806 | -0.94 |
| 26 | 2 | 1138.734965 | 1138.734471 | -0.43 |
| 27 | 2 | 1160.748073 | 1160.751550 | 3.00 |
| 28 | 2 | 1182.761180 | 1182.763850 | 2.26 |
| 29 | 2 | 1204.774288 | 1204.777197 | 2.41 |
| 30 | 2 | 1226.787395 | 1226.789992 | 2.12 |
| 31 | 2 | 1248.800503 | 1248.801050 | 0.44 |
| 32 | 2 | 1270.813610 | 1270.815971 | 1.86 |
| 33 | 2 | 1292.826718 | 1292.828203 | 1.15 |
| 34 | 2 | 1314.839825 | 1314.847791 | 6.06 |
| **Unassigned** | | | | |
| ? | 2 | | 606.902506 | |
| ? | 2 | | 628.917831 | |

Table C.2.: IRMPD-fragments assigned to the extraction of the horizontal scan of $TPGS_{18}{}^{+}$.

| **Fragment ion** | **Charge** | **Theoretical m/z** | **Observed m/z** | **Error [ppm]** |
|---|---|---|---|---|
| $f_{15}$ | 1 | 683.382445 | 683.381745 | -1.02 |
| $f_{16}$ | 1 | 727.408660 | 727.406864 | -2.47 |
| $e_{16}$ | 1 | 743.403575 | 743.401573 | -2.69 |
| ? | 1 | | 755.380561 | |

| | | | | |
|---|---|---|---|---|
| ? | 1 | | 769.419197 | |
| $f_{17}$ | 1 | 771.434875 | 771.435361 | 0.63 |
| $b_{15}$ | 1 | 799.393405 | 799.405923 | 15.66 |
| ? | 1 | | 813.446023 | |
| $f_{18}$ | 1 | 815.461090 | 815.462016 | 1.14 |
| $e_{18}$ | 1 | 831.456005 | 831.452934 | -3.69 |
| $b_{16}$ | 1 | 843.419620 | 843.432396 | 15.15 |
| $d_{18}$ | 1 | 887.482220 | 887.487126 | 5.53 |
| $c_{18}$ | 1 | 915.477135 | 915.478150 | 1.11 |
| $b_{18}$ | 1 | 931.472050 | 931.469868 | -2.34 |
| $a_{18}$ | 1 | 1079.560865 | 1079.560405 | -0.43 |
| TPGS18 | 1 | 1345.858215 | 1345.856405 | -1.34 |

Table C.3.: IRMPD-fragments assigned to the extraction of the horizontal scan of $TPGS_{22}{}^{2+}$. Ions marked with a * have been used to calibrate the entire 2D mass spectrum in the horizontal dimension.

| **Fragment ion** | **Charge** | **Theoretical m/z** | **Observed m/z** | **Error [ppm]** |
|---|---|---|---|---|
| $f_{10}$ | 1 | 463.25137 | 463.25168 | 0.68 |
| $f_{22}$* | 2 | 507.27759 | 507.27781 | 0.44 |
| $e_{22}$ | 2 | 515.27504 | 515.27477 | -0.52 |
| $g_{12}$ | 2 | 523.27250 | 523.27368 | 2.26 |
| $d_{22}$ | 2 | 543.28815 | 543.28802 | -0.24 |
| $c_{22}$ | 2 | 557.28561 | 557.28440 | -2.16 |
| $b_{22}$ | 2 | 565.28307 | 565.28252 | -0.97 |
| $g_{13}$ | 1 | 567.29872 | 567.29752 | -2.11 |
| $e_{13}$ | 1 | 611.32493 | 611.32410 | -1.35 |
| $a_{22}$* | 2 | 639.32747 | 639.32764 | 0.26 |
| $e_{14}$ | 1 | 655.35115 | 655.35013 | -1.55 |
| $f_{15}$ | 1 | 683.38245 | 683.38075 | -2.48 |
| $g_{16}$ | 1 | 699.37736 | 699.37609 | -1.81 |
| $g_{17}$ | 1 | 743.40358 | 743.40312 | -0.61 |
| $TPGS_{22}$* | 2 | 772.47615 | 772.47613 | -0.03 |
| $g_{19}$ | 1 | 831.45601 | 831.45425 | -2.11 |
| ? | 1 | | 843.43124 | |
| $f_{19}$ | 1 | 859.48731 | 859.48525 | -2.39 |
| $g_{20}$ | 1 | 875.48222 | 875.48079 | -1.64 |
| $f_{20}$ | 1 | 903.51352 | 903.51123 | -2.54 |
| $g_{21}$ | 1 | 919.50844 | 919.50274 | -6.20 |

| | | | | |
|---|---|---|---|---|
| $f_{21}$ | 1 | 947.53974 | 947.53839 | -1.42 |
| $f_{22}$* | 1 | 991.56595 | 991.56585 | -0.10 |
| $b_{20}$ | 1 | 1019.52448 | 1019.53987 | 15.09 |
| $d_{22}$ | 1 | 1063.58708 | 1063.58869 | 1.51 |
| $c_{22}$ | 1 | 1091.58200 | 1091.58381 | 1.66 |
| $b_{22}$ | 1 | 1107.57691 | 1107.57650 | -0.37 |
| $a_{22}$* | 1 | 1255.66573 | 1255.66748 | 1.40 |
| $TPGS_{22}$* | 1 | 1521.96308 | 1521.96240 | -0.45 |

Table C.4.: IRMPD-fragments assigned to the extraction of the horizontal scan of $TPGS_{28}{}^{3+}$.

| **Fragment ion** | **Charge** | **Theoretical m/z** | **Observed m/z** | **Error [ppm]** |
|---|---|---|---|---|
| $a_{17}$ | 1 | 1035.534650 | 1035.544766 | 9.77 |
| $TPGS_{17}$ | 1 | 1301.832000 | 1301.836254 | 3.27 |
| $a_{28}$ | 1 | 1519.823015 | 1519.820793 | -1.46 |
| $TPGS_{28}$ | 1 | 1786.120365 | 1786.088792 | -17.68 |
| $a_{28}$ | 2 | 771.406118 | 771.406320 | 0.26 |
| $a_{33}$ | 2 | 881.471655 | 881.471642 | -0.01 |
| $TPGS_{28}$ | 2 | 904.554793 | 904.542788 | -13.27 |

Table C.5.: IRMPD-fragments assigned to the extraction of the horizontal scan of Di-$TPGS_{21}{}^{2+}$.

| **Fragment Ion** | **Charge** | **Theoretical m/z** | **Observed m/z** | **Error [ppm]** |
|---|---|---|---|---|
| $f_{21}f$ | 2 | 498.272303 | 498.272065 | -0.48 |
| $a_{21}f$ | 2 | 630.322190 | 630.321768 | -0.67 |
| ? | 1 | | 685.438424 | |
| $a_{21}\,a$ | 2 | 762.372078 | 762.371755 | -0.42 |
| $f_{21}$ | 2 | 763.470865 | 763.471717 | 1.12 |
| ? | 2 | | 783.015076 | |
| $d_{21}$ | 2 | 799.481430 | 799.479768 | -2.08 |
| $c_{21}$ | 2.00 | 813.478888 | 813.48 | 2.23 |
| $a_{21}$ | 2 | 895.520753 | 895.521904 | 1.29 |
| $a_{21}f$ | 1 | 1237.655160 | 1237.658873 | 3.00 |
| $a_{21}\,a$ | 1 | 1501.754935 | 1501.759275 | 2.89 |
| $f_{21}$ | 1 | 1503.952510 | 1503.953557 | 0.70 |
| $e_{21}$ | 1 | 1521.963075 | 1521.968058 | 3.27 |
| $a_{21}$ | 1 | 1768.052285 | 1768.051816 | -0.27 |

| | | | | |
|---|---|---|---|---|
| Di-TPGS | 1 | 2034.349635 | 2034.349157 | -0.23 |

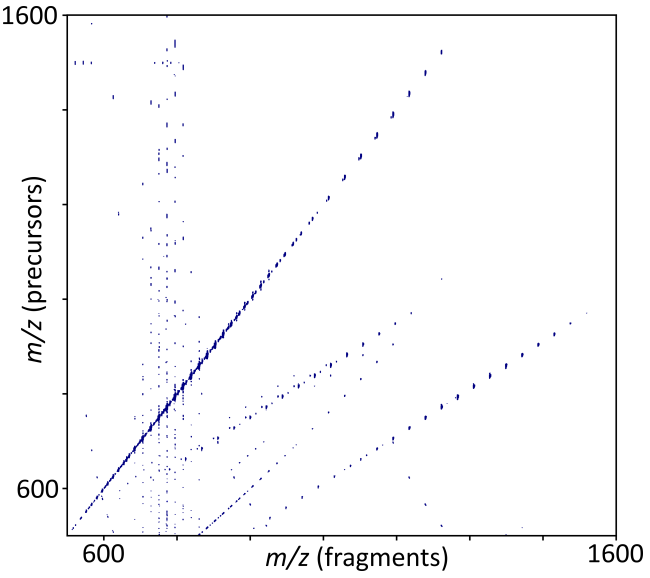



Figure C.4.: ECD 2D mass spectrum of TPGS in 100% acetonitrile and 1 mM $Na_2SO_4$. The autocorrelation line can clearly be observed, together with a diagonal line corresponding to loss of sodium, and curve harmonic lines. The vertical strikes of signal represent scintillation noise.

Table C.6.: ECD-fragments assigned to the extraction of the horizontal scan of $TPGS_{22}{}^{2+}$.

| **Fragment ion** | **Charge** | **Theoretical m/z** | **Observed m/z** | **Error [ppm]** |
|---|---|---|---|---|
| $TPGS_{22}$ | 2 | 772.476148 | 772.474210 | -2.51 |
| $f_{21}$ | 1 | 947.539735 | 947.542792 | 3.23 |
| ? | 1 | | 969.521499 | |
| $e_{21}$+Na | 1 | 987.532244 | 987.532466 | 0.23 |
| $f_{22}$+Na | 1 | 1015.563544 | 1015.556912 | -6.53 |
| $e_{22}$+Na | 1 | 1031.558459 | 1031.564040 | 5.41 |
| $d_{21}$+Na | 1 | 1043.558459 | 1043.559011 | 0.53 |
| ? | 1 | | 1069.542232 | |
| ? | 1 | | 1073.568741 | |
| $a_{22}$ | 1 | 1255.665725 | 1255.659458 | -4.99 |
| $TPGS_{22}$ | 1 | 1521.963075 | 1521.964921 | 1.21 |
| $TPGS_{21}$+2Na -H | 1 | 1499.918804 | 1499.922731 | 2.62 |

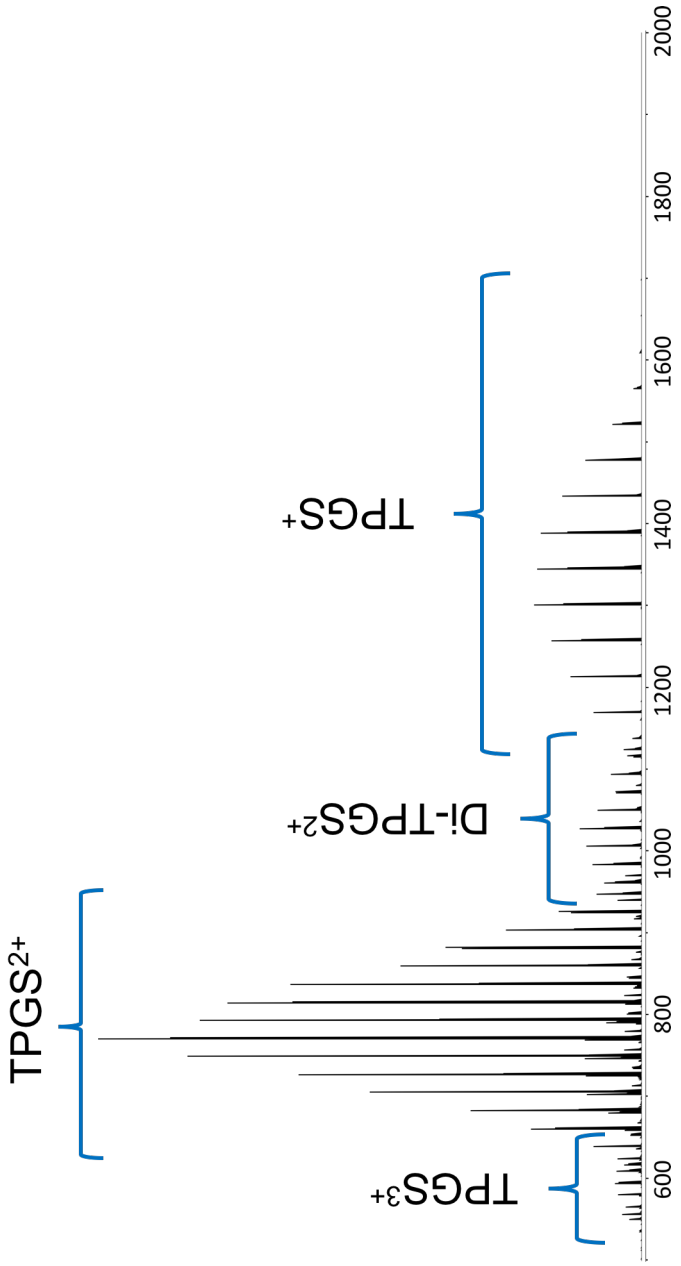



Figure C.1.: Full (1D) mass spectrum of TPGS in 100% acetonitrile and 1mM $Na_2SO_4$. The main polymeric distributions of TPGS are labelled. TPGS can be observed in three different charge distributions. Di-TPGS has been observed only with two positive charges. All the charges on the polymeric substance are given by sodium adducts.

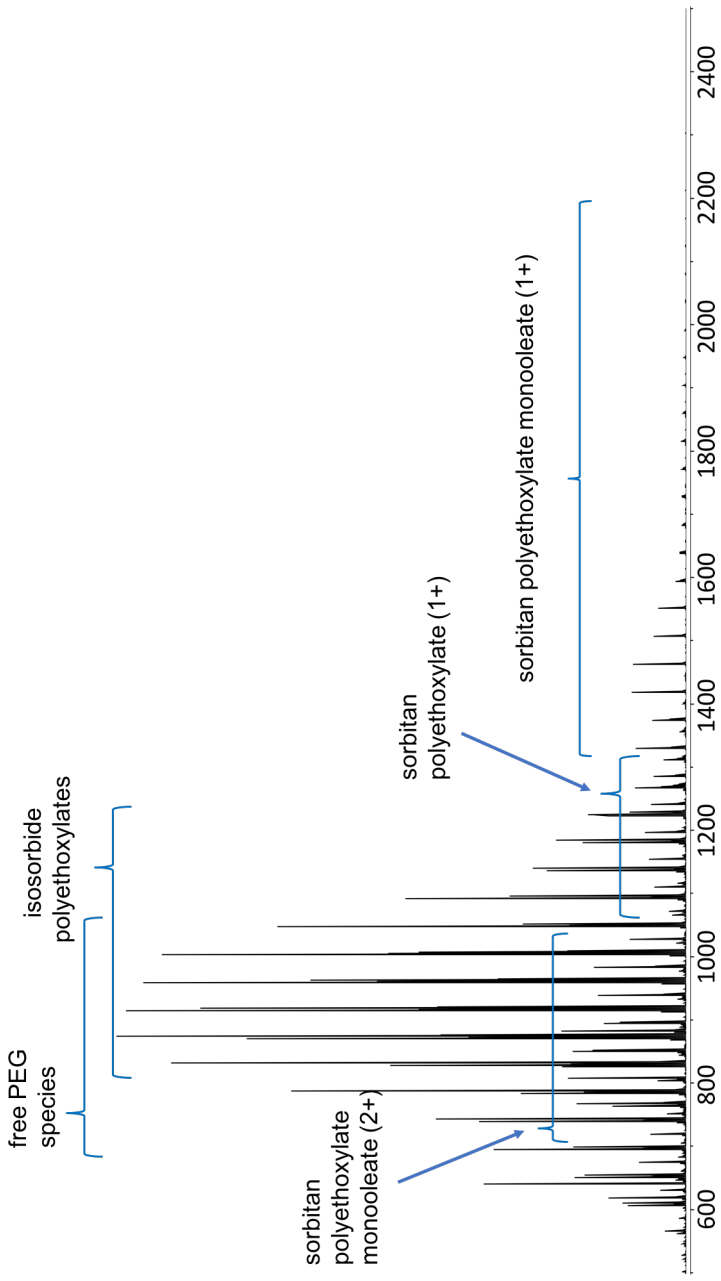



Figure C.2.: Full (1D) mass spectrum of Polysorbate 80 in 100% acetonitrile and 1mM $Na_2SO_4$. All the observed species were ionized by addiction of one or more sodium cations. The main species that could be detected are labelled in the figure, including sorbitan polyethoxilate and its monooleated version, both with one and two sodium adducts, free PEG, and monooleated and monolinoleated isosorbide polyethoxylates.

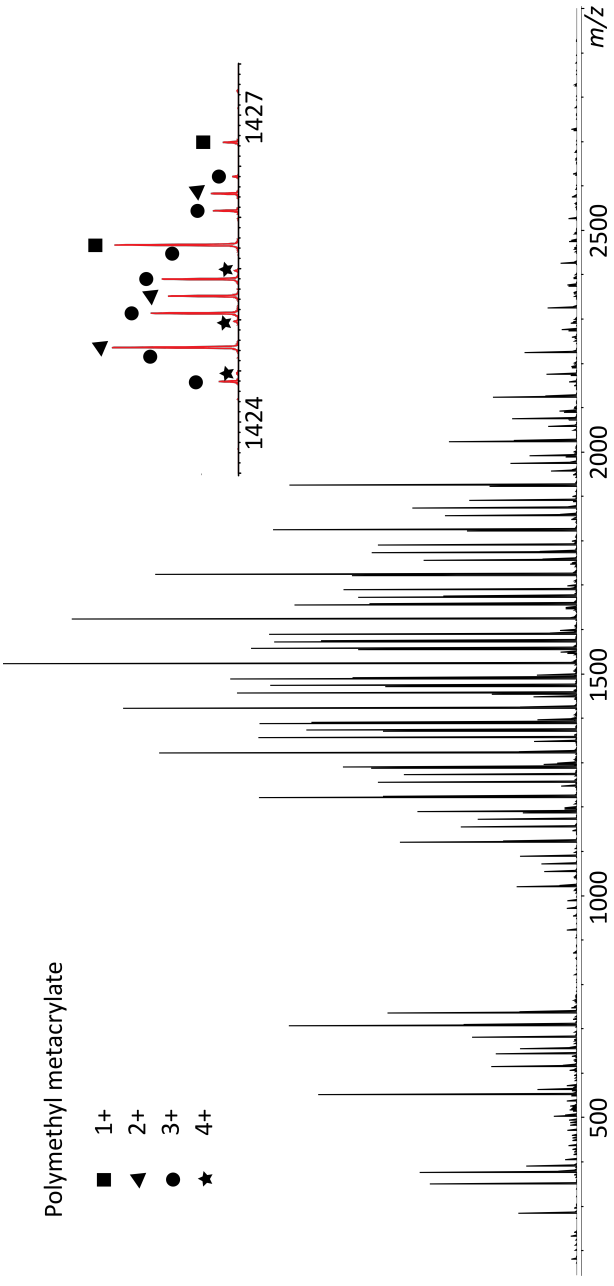



Figure C.3.: Full (1D) mass spectrum of PMMA 5 kDa in 100% acetonitrile and 1mM $Na_2SO_4$. Only distributions of PMMA with terminal H and sodium adducts were observed. The insert shows the complexity of the polymeric mixtures, with overlapping isotopic distributions of PMMA with different repeating units and charges.

Table C.7.: ECD-fragments assigned to the extraction of the horizontal scan of $TPGS_{27}{}^{3+}$.

| **Fragment ion** | **Charge** | **Theoretical m/z** | **Observed m/z** | **Error [ppm]** |
|---|---|---|---|---|
| $TPGS_{27}$+3$Na^+$ | 3 | 596.0241967 | 596.0222100 | -3.33 |
| $e_{27}$Na +H | 2 | 637.3396515 | 637.3409019 | 1.96 |
| $d_{26}Na^+2$ | 2 | 643.3393770 | 643.3389255 | -0.70 |
| $e_{14}Na^+H$ | 1 | 679.3487390 | 679.3495071 | 1.13 |
| $d_{13}$Na | 1 | 691.3487390 | 691.3488779 | 0.20 |
| $h_{22}$Na | 2 | 783.4671195 | 783.4592780 | -10.01 |
| ? | 2 | | 805.9638333 | |
| $h_{26}$Na | 2 | 871.5195495 | 871.5142626 | -6.07 |
| ? | 1 | | 1229.6129432 | |
| $e_{27}$Na | 1 | 1251.6895340 | 1251.6915927 | 1.64 |
| $TPGS_{27}$+$Na^+$ | 1 | 1742.0941500 | 1742.1492315 | 31.62 |

Table C.8.: ECD-fragments assigned to the extraction of the horizontal scan of Di-$TPGS_{21}{}^{2+}$.

| **Fragment ion** | **Charge** | **Theoretical m/z** | **Observed m/z** | **Error [ppm]** |
|---|---|---|---|---|
| Di-TPGS21 | 2 | 1028.669428 | 1028.668380 | -1.02 |
| $c_{21}\,e$Na | 1 | 1113.563939 | 1113.565850 | 1.72 |
| $c_{21}\,d$Na | 1 | 1169.590154 | 1169.592596 | 2.09 |
| $f_{21}$Na | 1 | 1526.942279 | 1526.957199 | 9.77 |
| $d_{21}$ | 1 | 1576.981465 | 1576.973195 | -5.24 |
| $d_{21}$Na | 1 | 1598.963409 | 1598.961621 | -1.12 |
| Di-$TPGS_{21}$ | 1 | 2034.349635 | 2034.350450 | 0.40 |

## C.4. 2D MS of Polysorbate80

Table C.9.: Peaks assigned to the autocorrelation line of the IRMPD 2D MS experiment of Polysorbate 80.

| Precursor ion | Charge | Theoretical m/z | Observed m/z | Error [ppm] |
|---|---|---|---|---|
| **Free Poly(ethylen glycol)** | | | | |
| 11 | 1 | 525.288150 | 525.288413 | 0.50 |
| 12 | 1 | 569.314365 | 569.314363 | 0.00 |
| 13 | 1 | 613.340580 | 613.339846 | -1.20 |
| 14 | 1 | 657.366795 | 657.366803 | 0.01 |
| 15 | 1 | 701.393010 | 701.393389 | 0.54 |
| 16 | 1 | 745.419225 | 745.420214 | 1.33 |
| **Poly(ethylen glycol) monolinoleate** | | | | |
| 13 | 1 | 875.570245 | 875.574238 | 4.56 |
| 14 | 1 | 919.596460 | 919.604258 | 8.48 |
| 15 | 1 | 963.622675 | 963.631109 | 8.75 |
| 16 | 1 | 1007.648890 | 1007.653662 | 4.74 |
| 17 | 1 | 1051.675105 | 1051.661872 | -12.58 |
| 20 | 1 | 1183.753750 | 1183.747603 | -5.19 |
| **Poly(ethylen glycol) monooleate** | | | | |
| 10 | 1 | 745.507250 | 745.507398 | 0.20 |
| 11 | 1 | 789.533465 | 789.532301 | -1.47 |
| 12 | 1 | 833.559680 | 833.558417 | -1.52 |
| 13 | 1 | 877.585895 | 877.586288 | 0.45 |
| 14 | 1 | 921.612110 | 921.615648 | 3.84 |
| 15 | 1 | 965.638325 | 965.639266 | 0.97 |
| 16 | 1 | 1009.664540 | 1009.667991 | 3.42 |
| **Isosorbide polyethoxylate** | | | | |
| 9 | 1 | 565.283065 | 565.283136 | 0.13 |
| 10 | 1 | 609.309280 | 609.308228 | -1.73 |
| 11 | 1 | 653.335495 | 653.334828 | -1.02 |
| 12 | 1 | 697.361710 | 697.362073 | 0.52 |
| 13 | 1 | 741.387925 | 741.388035 | 0.15 |
| 14 | 1 | 785.414140 | 785.415711 | 2.00 |
| 15 | 1 | 829.440355 | 829.440659 | 0.37 |
| 16 | 1 | 873.466570 | 873.482491 | 18.23 |
| 17 | 1 | 917.492785 | 917.501140 | 9.11 |
| **Isosorbide polyethoxylate monolinoleate** | | | | |
| 10 | 1 | 871.538945 | 871.541278 | 2.68 |

| | | | | |
|---|---|---|---|---|
| 11 | 1 | 915.565160 | 915.567865 | 2.95 |
| 12 | 1 | 959.591375 | 959.592951 | 1.64 |
| 13 | 1 | 1003.617590 | 1003.623834 | 6.22 |
| 14 | 1 | 1047.643805 | 1047.638627 | -4.94 |
| 15 | 1 | 1091.670020 | 1091.667442 | -2.36 |
| **Isosorbide polyethoxylate monooleate** | | | | |
| 7 | 1 | 741.475950 | 741.473149 | -3.78 |
| 8 | 1 | 785.502165 | 785.499859 | -2.94 |
| 9 | 1 | 829.528380 | 829.527846 | -0.64 |
| 10 | 1 | 873.554595 | 873.555602 | 1.15 |
| 11 | 1 | 917.580810 | 917.579193 | -1.76 |
| 12 | 1 | 961.607025 | 961.611428 | 4.58 |
| 13 | 1 | 1005.633240 | 1005.637076 | 3.81 |
| 14 | 1 | 1049.659455 | 1049.653122 | -6.03 |
| 15 | 1 | 1093.685670 | 1093.687256 | 1.45 |
| 16 | 1 | 1137.711885 | 1137.716520 | 4.07 |
| 17 | 1 | 1181.738100 | 1181.749242 | 9.43 |
| 18 | 1 | 1225.764315 | 1225.764798 | 0.39 |
| **Sorbitan polyethoxylate** | | | | |
| 24 | 2 | 633.338038 | 633.338667 | 0.99 |
| 25 | 2 | 655.351145 | 655.348228 | -4.45 |
| 11 | 1 | 671.346060 | 671.344852 | -1.80 |
| 26 | 2 | 677.364253 | 677.363278 | -1.44 |
| 27 | 1 | 699.377360 | 699.374738 | -3.75 |
| 28 | 2 | 721.390468 | 721.391952 | 2.06 |
| 29 | 2 | 743.403575 | 743.403822 | 0.33 |
| 13 | 1 | 759.398490 | 759.408201 | 12.79 |
| 31 | 2 | 765.416683 | 765.430365 | 17.88 |
| 32 | 2 | 787.429790 | 787.437998 | 10.42 |
| 19 | 1 | 1023.555780 | 1023.554529 | -1.22 |
| 20 | 1 | 1067.581995 | 1067.585841 | 3.60 |
| 21 | 1 | 1111.608210 | 1111.606016 | -1.97 |
| 22 | 1 | 1155.634425 | 1155.643785 | 8.10 |
| 23 | 1 | 1199.660640 | 1199.671470 | 9.03 |
| 24 | 1 | 1243.686855 | 1243.673206 | -10.97 |
| 25 | 1 | 1287.713070 | 1287.682036 | -24.10 |
| **Sorbitan polyethoxylate monooleate** | | | | |
| 26 | 2 | 809.486910 | 809.488632 | 2.13 |
| 28 | 2 | 853.513125 | 853.516155 | 3.55 |

| | | | | |
|---|---|---|---|---|
| 30 | 2 | 897.539340 | 897.542032 | 3.00 |
| 31 | 2 | 919.552448 | 919.536653 | -17.18 |
| 33 | 2 | 963.578663 | 963.568930 | -10.10 |
| 19 | 1 | 1287.801095 | 1287.821152 | 15.57 |
| 20 | 1 | 1331.827310 | 1331.839304 | 9.01 |
| 21 | 1 | 1375.853525 | 1375.863980 | 7.60 |
| 22 | 1 | 1419.879740 | 1419.883281 | 2.49 |
| 24 | 1 | 1507.932170 | 1507.916242 | -10.56 |
| 26 | 1 | 1595.984600 | 1595.977618 | -4.37 |
| 30 | 1 | 1772.089460 | 1772.100093 | 6.00 |
| 36 | 1 | 2036.246750 | 2036.244822 | -0.95 |

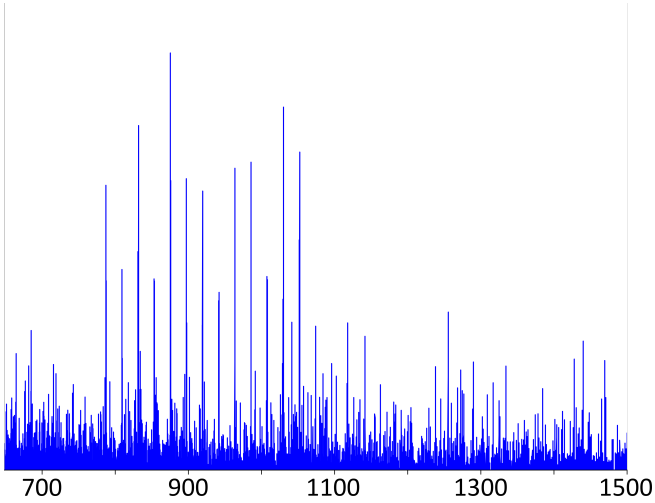



Figure C.5.: Neutral-loss line of all the doubly-charged precursors that lost an oleic acid fragment during IRMPD fragmentation period in the ICR-cell. Extracted from the IRMPD 2D mass spectrum of Figure 5.3 (left). All the peaks correspond to sorbitan polyethoxylate monooleate with two sodium adducts.

Table C.10.: IRMPD-fragments assigned to the extraction of the horizontal scan of $PEG_{12}$ monooleate.

| **Fragment ion** | **Suggested fragment** | **Charge** | **Theoretical m/z** | **Observed m/z** | **Error [ppm]** |
|---|---|---|---|---|---|
| $C_{16}H_{34}O_8Na$ | $B_8$ | 1 | 375.198940 | 375.200856 | 5.11 |
| $C_{16}H_{34}O_9Na$ | $A_8$ | 1 | 391.193855 | 391.196722 | 7.33 |
| $C_{18}H_{38}O_9Na$ | $B_9$ | 1 | 419.225155 | 419.226186 | 2.46 |
| $C_{18}H_{38}O_{10}Na$ | $A_9$ | 1 | 435.220070 | 435.221152 | 2.49 |
| $C_{20}H_{42}O_{10}Na$ | $B_{10}$ | 1 | 463.251370 | 463.251202 | -0.36 |

| | | | | | |
|---|---|---|---|---|---|
| $C_{20}H_{42}O_{11}Na$ | $A_{10}$ | 1 | 479.246285 | 479.245753 | -1.11 |
| $C_{22}H_{46}O_{11}Na$ | $B_{11}$ | 1 | 507.277585 | 507.277125 | -0.91 |
| $C_{24}H_{50}O_{12}Na$ | $B_{12}$ | 1 | 551.303800 | 551.302212 | -2.88 |
| $C_{27}H_{53}O_{14}Na$ | | 1 | 623.324930 | 623.323248 | -2.70 |
| $C_{28}H_{54}O_{14}Na$ | | 1 | 637.340580 | 637.340026 | -0.87 |
| $C_{31}H_{60}O_{14}Na$ | | 1 | 679.387530 | 679.391965 | 6.53 |
| $C_{35}H_{68}O_{14}Na$ | | 1 | 735.450130 | 735.449144 | -1.34 |
| PEG monoleate ($C_{42}H_{82}O_{14}Na$) | $[M+Na^+]^+$ | 1 | 833.559680 | 833.559098 | -0.70 |

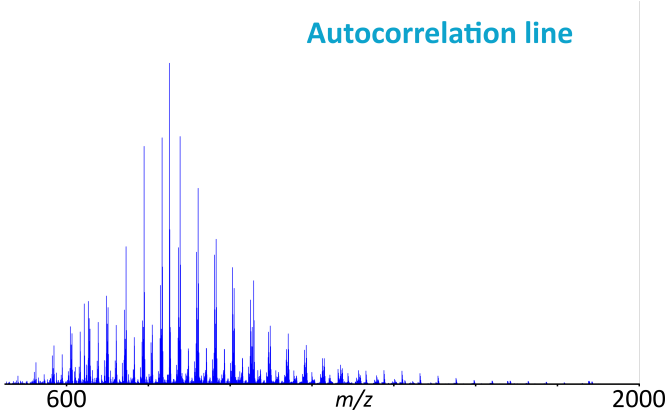



Figure C.6.: Autocorrelation line extracted from the ECD 2D mass spectrum of Figure 5.3.a. A difference in ion intensities for all the species can be observed compared to the 2D IRMPD mass spectrum, due to the different dissociation technique.

Table C.11.: ECD-fragments assigned to the extraction of the horizontal scan of $[(PEG)_{32}$ sorbitan polyethoxylate monooleate$]^{2+}$.

| **Fragment ion** | **Suggested fragment** | **Charge** | **Theoretical m/z** | **Observed m/z** | **Error [ppm]** |
|---|---|---|---|---|---|
| Polysorbate monooleate | $[M+2Na^+]$ | 2 | 941.565555 | 941.565548 | -0.01 |
| $C_6H_{10}O_4(C_2H_4O)_{29}Na$ | $B_{29}$ | 1 | 1427.796800 | 1427.760905 | -25.14 |
| $C_6H_{10}O_4(C_2H_4O)_{31}Na$ | $B_{31}$ | 1 | 1533.859795 | 1533.859972 | 0.12 |
| $C_6H_9O_4(C_2H_4O)_{31}Na_2$ | $B_{31}$ | 1 | 1555.841739 | 1555.840127 | -1.04 |
| $C_6H_{11}O_5(C_2H_4O)_{32}Na_2$ | Oleic fragment loss | 1 | 1617.878519 | 1617.878311 | -0.13 |
| Polysorbate monooleate | Na loss | 1 | 1860.141890 | 1860.141930 | 0.02 |

## C.5. 2D MS of PMMA

Table C.12.: ECD-fragments assigned to the extraction of the horizontal scan of $[(MMA)_{40}Na_3]^{3+}$.

| Fragment ion | Charge | Theoretical m/z | Observed m/z | Error [ppm] |
|---|---|---|---|---|
| $PMMA_{40}$ | 3 | 1357.693503 | 1357.692809 | -0.51 |
| $PMMA_{33}$ | 2 | 1674.862140 | 1674.865544 | 2.03 |
| $PMMA_{37}$ methyl loss | 2 | 1867.455263 | 1867.476767 | 11.52 |
| $PMMA_{37}$ | 2 | 1874.967000 | 1874.962203 | -2.56 |
| Methyl loss, sodium loss, H gain? | 2 | 2018.037820 | 2018.037064 | -0.37 |
| $PMMA_{40}$ methyl loss | 2 | 2029.028792 | 2029.030880 | 1.03 |

Table C.13.: ECD-fragments assigned to the extraction of the horizontal scan of $[(MMA)_{42}Na_4]^{4+}$.

| Fragment ion | Charge | Theoretical m/z | Observed m/z | Error [ppm] |
|---|---|---|---|---|
| $PMMA_{42}$ | 4 | 1074.043648 | 1073.995409 | -44.91 |
| $PMMA_{39}$ | 3 | 1324.342693 | 1324.323956 | -14.15 |
| $PMMA_{42}Na_3$-$CH_3$ | 3 | 1427.050555 | 1427.060566 | 7.02 |
| $PMMA_{37}$ | 2 | 1874.967000 | 1874.975471 | 4.52 |
| $PMMA21Na_2$-$CH_3$ | 1 | 2134.072194 | 2134.057659 | -6.81 |

Table C.14.: IRMPD-fragments assigned to the extraction of the horizontal scan of $[(MMA)_{42}Na_4]^{4+}$.

| Fragment ion | Charge | Theoretical m/z | Observed m/z | Error [ppm] |
|---|---|---|---|---|
| ? | 4 | | 837.375335 | |
| ? | 4 | | 837.502483 | |
| $PMMA_{31}$ | 3 | 1057.536213 | 1057.537709 | 1.41 |
| ? | 3 | | 1080.212129 | |
| $PMMA_{32}$ | 3 | 1090.887023 | 1090.887455 | 0.40 |
| $PMMA_{31}$ | 2 | 1574.809710 | 1574.799696 | -6.36 |
| $PMMA_{47}$ (adduct) | 3 | 1591.149173 | 1591.136386 | -8.04 |
| $PMMA_{37}$ | 2 | 1874.967000 | 1874.980840 | 7.38 |

| | | | | |
|---|---|---|---|---|
| $PMMA_{21}$ | 1 | 2126.105900 | 2126.100147 | -2.71 |

## C.6. Parameters for 2D MS acquisition

Table C.15.: Parameters for two-dimensional MS acquisition of the polymer samples.

| Parameters | TPGS | | Polysorbate 80 | | PMMA | |
|---|---|---|---|---|---|---|
| | IRMPD | ECD | IRMPD | ECD | IRMPD | ECD |
| **Acquisition mode** | | | | | | |
| Acquisition size [datapoints, 16 bit] | 512k | 512k | 512k | 256k | 512k | 512k |
| Horizontal low mass cut-off [m/z] | 368.2 | 368.2 | 368.2 | 368.2 | 368.2 | 368.2 |
| Horizontal high mass cut-off [m/z] | 3000 | 3000 | 3000 | 3000 | 3000 | 3000 |
| Transient length, $t_2$ [s] | 0.2097 | 0.2097 | 0.2097 | 0.1049 | 0.2097 | 0.2097 |
| Ion accumulation time [s] | 0.2 | 0.2 | 0.2 | 0.2 | 0.2 | 0.2 |
| Ion cooling time [s] | 0.1 | 0.1 | 0.1 | 0.1 | 0.1 | 0.1 |
| **Source (nESI)** | | | | | | |
| Capillary voltage [V] | -4800 | -4800 | -4800 | -4800 | -4800 | -4800 |
| Deflector plate [V] | 220 | 220 | 220 | 220 | 220 | 220 |
| Funnel [V] | 150 | 150 | 150 | 150 | 150 | 150 |
| Skimmer [V] | 10 | 10 | 10 | 10 | 10 | 10 |
| Funnel RF [Vpp] | 200 | 200 | 200 | 200 | 200 | 200 |
| **Collision cell** | | | | | | |
| Collision voltage [V] | -4 | -4 | -4 | -4 | -4 | -4 |
| DC extract bias [V] | 0.6 | 0.6 | 0.6 | 0.6 | 0.6 | 0.6 |

| | | | | | | |
|---|---|---|---|---|---|---|
| RF frequency [MHz] | 2 | 2 | 2 | 2 | 2 | 2 |
| Collision RF Amplitude [Vpp] | 1400 | 1400 | 1400 | 1400 | 1400 | 1400 |
| **Octopole** | | | | | | |
| Frequency [MHz] | 2 | 2 | 2 | 2 | 2 | 2 |
| RF Amplitude [Vpp] | 1400 | 1400 | 1400 | 1400 | 1400 | 1400 |
| **Transfer optics** | | | | | | |
| Time of Flight [s] | 0.001 | 0.001 | 0.001 | 0.001 | 0.001 | 0.001 |
| Frequency [MHz] | 4 | 4 | 4 | 4 | 4 | 4 |
| RF Amplitude [Vpp] | 350 | 350 | 350 | 350 | 350 | 350 |
| **ICR-cell** | | | | | | |
| Transfer exit Lens [V] | -15 | -15 | -15 | -15 | -15 | -15 |
| Analyser entrance [V] | -8 | -8 | -8 | -8 | -8 | -8 |
| Side kick [V] | 10 | 10 | 10 | 10 | 10 | 10 |
| Side kick offset [V] | -5 | -5 | -5 | -5 | -5 | -5 |
| Front Trap plate [V] | 0.55 | 0.6 | 0.6 | 0.6 | 0.6 | 0.6 |
| Back Trap plate [V] | 0.57 | 0.62 | 0.62 | 0.62 | 0.62 | 0.62 |
| Sweep Excitation power, P3 [%] | 12 | 12 | 12 | 12 | 12 | 12 |
| **2D MS pulse programme** | | | | | | |
| $t_1$ [s] | 1.00E-06 | 1.00E-06 | 1.00E-06 | 1.00E-06 | 1.00E-06 | 1.00E-06 |

| | | | | | | |
|---|---|---|---|---|---|---|
| $t_1$ increment [s] | 1.00E-06 | 1.00E-06 | 1.00E-06 | 1.00E-06 | 1.00E-06 | 1.00E-06 |
| $P_1$ and $P_2$ pulse length [s] | 4.00E-07 | 4.00E-06 | 4.00E-07 | 4.00E-06 | 4.00E-07 | 4.00E-06 |
| $P_1$ and $P_2$ excitation power [Db] | 10.42 | 10.42 | 10.42 | 10.42 | 10.42 | 10.42 |
| Loops | 2048 | 2048 | 4096 | 2048 | 4096 | 8192 |
| **In-cell Fragmentation** | | | | | | |
| Laser Power (25W) [%] | 50 | / | 50 | / | 50 | / |
| IR pulse length [s] | 0.34 | / | 0.28 | / | 0.5 | / |
| ECD pulselength [s] | / | 0.02 | / | 0.02 | / | 0.03 |
| ECD bias | / | 1.5 | / | 1.5 | / | 1.5 |
| ECD Extraction lens | / | 10 | / | 10 | / | 10 |

# D. Supporting Information for Chapter 6

## D.1. Detailed construction of the ICR-cell

Figure D.1.: Alingnment ring, with circular holes to embed the ceramic rods, and four threaded holes to accommodate the aluminium rods for installation into the home-built system. In correspondence of the red signs on the ring structure, teflon plugs can be seen on the outer part of the ring. The teflon plugs, positioned at 2.00 mm from the ring, allow alignment of the cell in the flight tube that will be inserted inside the magnet bore.

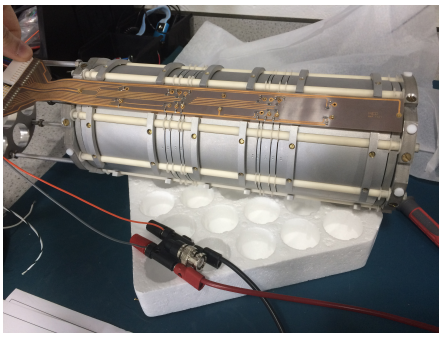


Figure D.3.: Testing of the PCB on the ICR-cell. A perfect fitting and good connections were confirmed.

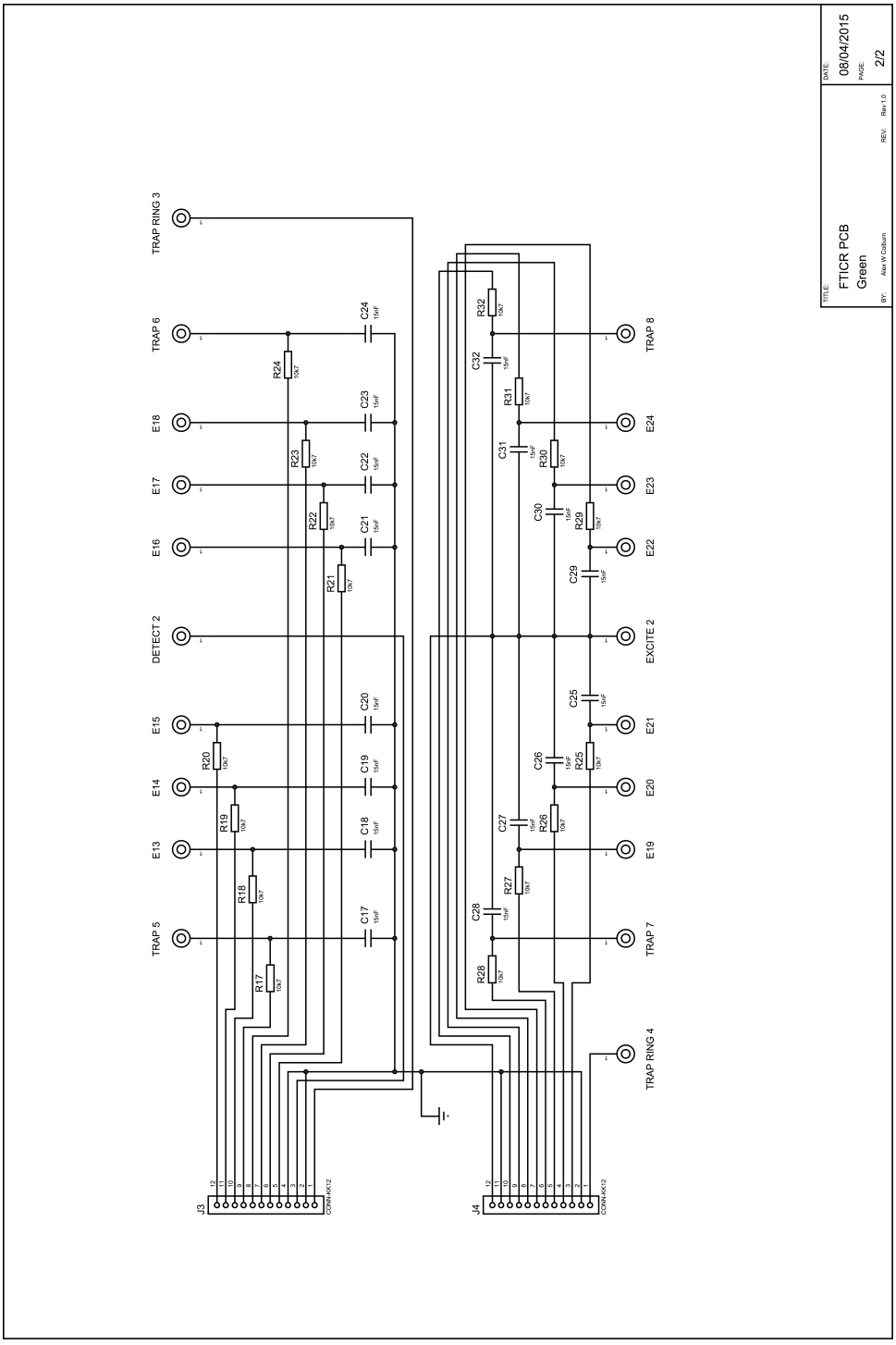



Figure D.2.: Electrical diagram of a printed circuit board.

Figure D.4.: Details of the final assembly of the ICR-cell. The four "sides" of the cell, aligned and clamped to the ceramic rods (left) and the inner trapping rings (right). Care is necessary when assembling the different parts, in order not to damage the ceramic rods and not to scratch the thin layer of gold/nickel on the plate surfaces.

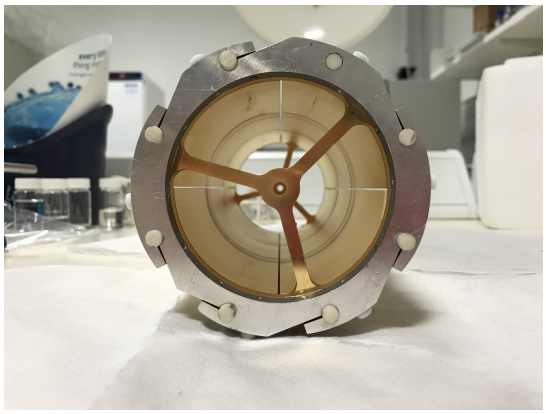


Figure D.5.: Inside the ICR-cell, the four quarters are perfectly assembled in a smooth-inner surface cylinder.

Figure D.6.: The assembled ICR-cell, with the aluminium structure for the installation inside the mass spectrometer. The outer structure rings have been implemented with teflon plugs to improve centering in the flight tube.